

Sorting and the output loss due to search frictions^{*}

Pieter A. Gautier[†] and Coen N. Teulings[‡]

Abstract

We analyze a general search model with on-the-job search (OJS) and sorting of heterogeneous workers into heterogeneous jobs. For given values of non-market time, the relative efficiency of OJS, and the amount of search frictions, we derive a simple relationship between the unemployment rate, mismatch and wage dispersion. We estimate the latter two from standard micro data. Our methodology accounts for measurement error, which is crucial to distinguish true from spurious mismatch and wage dispersion. We find that without frictions, output would be about 9.5%

^{*}We thank seminar participants at MIT, the 2009 Sandjberg conference on search models of the labor market, the Sciences Po conference on sorting, University of Munich, SED (2011), the 2011 Tinbergen conference in Amsterdam, Essex, UPenn and the SAM meeting in Rouen for useful comments and discussions. Finally, we thank Bart Hobijn for sharing his labor-market flow data with us, and Xiaoming Cai for excellent research assistance.

[†]VU University Amsterdam, Tinbergen Institute, CEPR, Ces-Ifo, IZA, email: pgautier@feweb.vu.nl

[‡]University of Cambridge, University of Amsterdam, Tinbergen Institute, CEPR, CeS-Ifo, IZA, email:

higher if firms *can* commit to pay wages as a function of match quality and 15.5% higher if they *cannot*. Non-commitment leads to a business-stealing externality which causes a 5.5% drop in output.

JEL codes: E24, J62, J63, J64

1 Introduction

Labor productivity depends crucially on the proper sorting of worker types into job types, a process that is hindered by search frictions. If all unemployed workers and jobs were alike, it would be hard to imagine why it takes workers months to find a suitable job.

But also if workers and jobs were heterogeneous but search frictions were absent, the loss in output due to mismatch would be irrelevant because all workers would be matched to their optimal job types. This paper explicitly models this interaction between search frictions and heterogeneity and estimates the output loss due to search frictions. This loss is then decomposed in its three components: (i) unemployment, (ii) resources spent on recruitment activities and (iii) mismatch. We also estimate the output loss that can be attributed to the inability of employers to commit to future wage payments. Only this

loss can potentially be reduced by policy intervention.

The starting point of this paper is the framework of Gautier, Teulings and Van Vuuren (2010) who analyze a class of search models with on-the-job (OJS) search and worker and job heterogeneity where the productivity of a match depends on the degree of mismatch. Their production function can be interpreted as a second-order Taylor approximation of a more general specification of the production technology. Within this framework, various wage mechanisms can be analyzed such as wage posting with full commitment, as in Burdett and Mortensen (1998) and Bontemps, van den Berg, and Robin (2000), and wage mechanisms without commitment, as in Coles (2001) and Shimer (2006). The key difference between wage setting with and without commitment is that in the former case, firms pay both hiring and no-quit premiums, whereas in the latter case, firms pay only

no-quit premiums. Our model is related to hedonic pricing/assignment/sorting models in which worker types are imperfect substitutes –in the spirit of Rosen (1974), Sattinger (1975) and Teulings (1995, 2005). Intuitively, the less substitutable worker types are, the larger will be the productivity loss due to mismatch. We use a relation between Katz and Murphy’s (1992) elasticity of complementarity between high- and low-skilled workers and the curvature of the production function. This curvature determines how sensitive output is to the degree of mismatch. In a Walrasian equilibrium, the above models generate a perfect sorting of high-skilled workers at complex jobs. With search frictions, this perfect correlation breaks down. Within the class of sorting models with search frictions, a distinction can be made between hierarchical models, like Shimer and Smith (2000), and circle models, like Marimon and Zilibotti (1999). Since circle models

are more easy to handle analytically than hierarchical models, we apply a circle model in our theoretical analysis.¹

We show that the equilibrium depends on just four parameters: (i) the value of non-market time, (ii) the relative efficiency of on- versus off-the-job search, (iii) the curvature of the production function (i.e. how fast output falls with the degree of mismatch), and (iv) a composite parameter that measures the amount of search frictions. The relevance of our model depends on how well it can match the empirical values for these parameters.

The main hurdle is to obtain estimates for wage dispersion and the output loss due to mismatch. We follow Gautier and Teulings (2006) by using data on wages and on worker

¹We can still match moments generated by a hierarchical process because analytical conclusions from a circle model translate into a hierarchical setting because the former can be viewed of as a second order Taylor approximation of the latter, see Gautier, Teulings, and Van Vuuren (2005). The intuition is that in hierarchical models, the output and wage loss due to mismatch also depends on the expected distance to one's optimal job type.

and job characteristics to construct an empirical measure for mismatch. We also offer a new simple statistic for wage dispersion due to search frictions, namely the intercept of a simple quadratic wage regression with appropriately normalized measures for worker and job characteristics (this measures the difference between the expected maximum and the average wage). The simplicity of this measure makes it easily applicable. Since observed mismatch can be either real or due to measurement error, it is important to correct for this. We show how to do that. Given wage dispersion and mismatch, our model implies a value for the unemployment rate. We find values that are close to the empirical values of about 5%. Hence, our model can jointly explain the observed wage dispersion for workers with equal skill and unemployment.

We use the model to calculate the total output loss due to search frictions, which

we estimate to be between 9 and 16% of the total value added of labour, depending on whether firms can or cannot commit to paying hiring premiums. Unemployment accounts for less than 30% of this loss. If firms cannot commit to wages, their quasi-rents are higher than in the social optimum, due to a business-stealing externality. As a result of free entry, these quasi-rents are all spent on (excess) vacancy creation. The estimated output loss due to this business-stealing externality is 5.5% of the total value added of labour. This externality can potentially be reduced by policies that shift rents from the firms to the workers.²

Most of the literature on sorting with frictions considers global absolute advantages of high-skilled workers. Atakan (2006) and Eeckhout and Kircher (2009) consider a simplified

²In Gautier et al. (2014) we show that if a union sets pay scales ex ante to reduce the business externality, it will set the lowest wage on the pay scale too high which increases unemployment and may actually reduce welfare.

version of Shimer and Smith (2000). Similar to our model, wages are highest at the optimal assignment and they are lower at both less and more complex jobs. Their model is however less suitable to bring to the data. Hagedorn, Law and Manovskii (2012) show how both the Shimer and Smith (2000) model and our model are non-parametrically identified from individual wage data from firm-worker data sets. The idea is that wages are informative about the ranks of worker types within each firm. Bagger and Lentz (2012) consider a sorting model where workers search most intensively for the jobs where they earn the highest wages. Lise and Robin (2013) consider a sorting model with on-the-job search that focuses on the macro dynamics in the presence of aggregate shocks. Lise, Meghir and Robin (2012), Lopes de Melo (2008) and Bartolucci and Devicienti (2012) also look at sorting in models with OJS. Their focus is on interpreting the correlations

between worker and firm fixed effects. Under comparative advantage, this issue is not very meaningful. For example, in Teulings and Gautier (2004), complex jobs do not have an absolute advantage over simple jobs. They only have a comparative advantage when occupied by better-skilled workers. Since skilled workers have an absolute advantage, workers employed in more-complex jobs earn higher wages but the higher wages are not due to the job, but to the type of workers that occupy these jobs. In this context, one can just reverse the ordering of job types to change from negative to positive assortative matching. Without loss of generality, we focus on the latter. An important difference between the empirical sorting models described above and ours is that all jobs in a firm have the same fixed effect. We do not make that assumption. Cornfeld (2014) considers a different type of sorting model where skill is defined as the set of tasks that a worker

can perform. Finally, Jovanovic (2013) looks at the effect of misallocation in the labor market on growth in a model where agents must learn about their abilities and

Concerning wage dispersion, Hornstein et al. (2010) also derive a simple relationship between the unemployment rate and wage dispersion, the *mean-min* ratio. We show that this measure is sensitive to measurement error. They argue that search models without OJS cannot explain the coexistence of a low unemployment rate and substantial wage dispersion because the former suggests low frictions, while the latter suggests high frictions. Gautier and Teulings (2006) made a similar point. This issue can be resolved by allowing for OJS, since this lowers the reservation wage (consequently wage dispersion rises and the unemployment rate falls). Allowing for OJS is also quantitatively important, since Fallick and Fleischman (2004) and Nagypal (2005) show that job-to-job flows are

substantial.

The rest of this paper is organized as follows. Section 2 presents a summary of the model of Gautier, Teulings and Van Vuuren (2010) as a point of reference for the rest of the paper. Section 3 discusses how we can identify and measure mismatch and wage dispersion in the presence of measurement error. Section 4 presents the calibration results, the estimation of the output loss due to search frictions and the decomposition of this loss. Finally, Section 5 concludes.

2 The Model

2.1 Assumptions

Production

There is a continuum of worker types, s , and job types, c ; s and c are locations on a circle.

Workers can only produce output when matched to a job. The productivity of a match of worker type s to job type c depends on the shortest distance $|x|$ between s and c along the circumference of the circle. $Y(x)$ has an interior maximum at $x = 0$ and is symmetric around this maximum $Y(0)$ (normalized to unity). Finally, $Y(x)$ is twice differentiable and strictly concave. We consider the simplest functional form that meets these criteria:

$$Y(x) = 1 - \frac{1}{2}\gamma x^2. \tag{1}$$

We call x the *mismatch* indicator. The parameter γ determines the substitutability of worker types: the lower γ , the more easily worker types can be substituted. $Y(x)$ can be interpreted as a second-order Taylor approximation around the optimal assignment of a more general production technology. Since the first derivative of a continuous production function equals zero in the optimal assignment, $Y'(0) = 0$, the first-order term drops out.

We are interested in equilibria where unemployed job seekers do not accept all job offers,

which imposes a minimum constraint on γ .³

Labor supply and the value of non-market time

Labor supply per s -type is uniformly distributed over the circumference of the circle.

Total labor supply in period t equals $L(t)$. We normalize the labor force at $t = 0$ to

one. Unemployed workers receive the value of non-market time B . Employed workers

supply a fixed amount of labor (normalized to one), and their payoff is equal to the wage

they receive. Workers live forever. They maximize the discounted value of their expected

lifetime payoffs.

Golden-growth path

³A sufficient condition for this is that $Y(x) < 0$ for at least some x .

We study the economy while it is on a golden-growth path, where the discount rate $\rho > 0$ is equal to the growth rate of the labor force. Hence, the size of the labor force is $L(t) = \exp(\rho t)$. The assumption of a golden-growth path buys us a lot in terms of transparency and tractability. The golden-growth assumption is a generalisation of the assumption of zero discounting (zero discounting is the special case of the golden-growth being equal to zero), an assumption that is often applied in the wage posting literature, see for example Burdett and Mortensen (1998). New workers enter the labor force as unemployed. Since labor supply at $t = 0$ and the productivity in the optimal assignment $Y(0)$ are normalized to one, the output of this economy would be equal to one in the absence of search frictions.

Job offer arrival rates and job destruction

Unemployed job seekers receive job offers at a rate λ . Workers receive job offers at a rate $\psi\lambda$. The parameter $\psi, 0 \leq \psi \leq 1$, measures the efficiency of on- relative to off-the-job search; $\psi = 0$ is the case without OJS; $\psi = 1$ is the case where on- and off-the-job search are equally efficient. Matches between workers and jobs are destroyed at an exogenous rate $\delta > 0$.

As is well known in the job search literature, see e.g. Burdett and Mortensen (1998), the number of parameters can be reduced by introducing a composite parameter⁴,

$$\kappa \equiv \frac{2\lambda}{\rho + \delta}.$$

Hence, we can ignore the separate parameters ρ, δ , and λ , and focus on the composite parameter κ instead.

⁴It is convenient to add a factor 2 to the definition of κ to account for the fact that this model is symmetric around the optimal allocation $x = 0$. Hence, job offers with positive and negative values of x are equivalent.

Vacancy creation and contact technology

For our empirical analysis, we can ignore the process of vacancy creation and the job offer arrival technology that underlies the value of κ . However, when analyzing the output loss due to search frictions and the constrained efficiency of the equilibrium, we have to be explicit about vacancy creation and the contact technology. We assume that there is free entry of vacancies for all c -types. The cost of maintaining a vacancy is equal to K per period. After a vacancy is filled, the firm's only cost is the worker's wage. The supply of vacancies is determined by a zero profit condition. Vacancies are uniformly distributed over the circumference of the circle. When a worker leaves a job, this job disappears.

Let u be the unemployment rate. Due to the normalization of labor supply to one, u is equal to the number of unemployed. Then, the effective number of job seekers is

equal to the number of unemployed plus the number of employed weighted by the relative efficiency of on-the-job search, $u + \psi(1 - u)$. The job-offer-arrival rate λ is a function of the effective supply of job seekers and the number of vacancies, v :

$$\lambda = \lambda_0 [u + \psi(1 - u)]^{-\beta} v^\alpha, \quad (2)$$

where $0 \leq (\alpha, \beta) \leq 1$. This specification embodies two important special cases: (i) for

$\psi = 0$ and $\beta = \alpha, 0 < \alpha < 1$, $\lambda = \lambda_0 (u/v)^{-\alpha}$: the classical Pissarides constant-returns-

to-scale matching function; (ii) for $\alpha = 1, \beta = 0$: the quadratic contact technology. Note

that for $\psi = 1$, the value of β is irrelevant. Hence, the case $\alpha = 1, \psi = 1$ is equivalent to

the quadratic contact technology in this setting. Finally, note that for our purposes, we

do not need to know K because any decrease in K can be captured by a corresponding

increase in κ .^{5 6}

Wage setting

Wages, denoted by $W(x)$, are set unilaterally by the firm, conditional on the mismatch indicator x in the current job. We analyze wage setting under two different assumptions.

Under the first assumption, firms can commit to a future wage payment contingent on x .

Then, firms pay both no-quit and hiring premiums. That is, they account for the positive

effect of a higher wage offer on reduced quitting and increased hiring as in Burdett and

⁵Shifting K to $2K$ so that v shifts to $v/2$ together with shifting λ_0 to $2^\alpha \lambda_0^{-\alpha}$ does not effect the equilibrium.

⁶The number of free parameters can be reduced even further. If we replace γ by $\zeta \equiv \kappa^{-2}\gamma$, κ can be normalized to one. When we simultaneously increase κ to 2κ and γ to 4γ , this is equivalent to increasing simultaneously the job search efficiency and the cost of a bad match. As a result, the upper bound \bar{x} would shrink to $\frac{1}{2}\bar{x}$, but for the rest, everything would remain the same. The composite parameter ζ can be interpreted as a summary statistic for search frictions. Details of this transformation are in Web Appendix C.4 In our empirical application, we need a particular normalization for x and, so we do not apply the final normalization here.

Mortensen (1988). Under the second assumption, firms are unable to commit to future wage payments. In this case, hiring premiums are non-credible because immediately after the worker has accepted the job, the firm has no incentive to continue paying a hiring premium, since the worker cannot return to her previous job. Workers anticipate this, and will therefore not respond to this premium in the first place, which means that firms will not offer it. No-quit premiums are credible even without commitment because it is in the firm's interest to pay them as soon as the worker has accepted the job, for if the firm does not pay them the worker will quit as soon as a better outside offer arrives.

Since the equilibrium of this economy and its comparative statics are analyzed extensively in Gautier, Teulings, and Van Vuuren (2010), we will only provide a short summary of the main results that are needed for the empirical implementation, below.

2.2 Characterization of the equilibrium

The equilibrium of this economy is characterized by a wage function $W(x)$ and an upper bound \bar{x} for the absolute value of the mismatch indicator $|x|$. Job offers with a higher value of x will not be accepted. Since the model is symmetric around $x = 0$, $W(|x|) = W(x)$.

For sake of notational convenience, we focus on the case $x \geq 0$. Wages are decreasing in the mismatch indicator x : the lower the mismatch, the higher the wage rate paid by firms.⁷ The upper bound \bar{x} implies a value for u (for a derivation see Web Appendix C.1)

$$u = \frac{1}{1 + \kappa \bar{x}}.$$

Note that the model is very similar to the stochastic job search model of Pissarides (2000)

extended with on-the-job search and the constraint that the derivative of $Y(x)$ is zero in

⁷See Gautier et al. (2010) for a proof. The logic is the same as why bid functions in auction theory are increasing in valuations.

the optimal assignment, $x = 0$.⁸ A worker accepts any job offer with a wage above his current wage and consequently with a mismatch indicator smaller than in his current job.

Unemployed workers accept only job offers with $x < \bar{x}$.

Bellman equations under the Golden Growth assumption

Due to the golden-growth assumption, asset values for job seekers and employed workers take a simple form that can be easily interpreted. Let V^U and V^E be the asset values of an unemployed and an employed worker at her marginal job type (with mismatch indicator \bar{x}) respectively and let $E_x Y$ and $E_x W$ denote the expected output and wage respectively

(the expectation being taken over the mismatch indicator x among employed workers).

⁸The only difference is that under free entry the composition of vacancies adjusts to the composition of the unemployment.

Then

$$\rho V^U = uB + (1 - u) E_x W, \quad (3)$$

$$\rho V^E = \frac{uW(\bar{x}) + \psi(1 - u) E_x W}{u + \psi(1 - u)},$$

$$vK = (1 - u) (E_x Y - E_x W).$$

The derivation of these Bellman equations can be found in Web Appendix C.2. The

values of unemployment and employment are a weighted average of the expected payoffs

in the states of employment and unemployment. For the value of unemployment, the

expected payoffs are weighted by their share in the total population. The total expected

cost of vacancy creation are equal to expected profits, which in turn are equal to expected

productivity minus expected wages times employment. The Bellman equations take this

simple form due to the Golden Growth assumption that the growth rate of the workforce

is equal to the discount rate.

The output loss due to search frictions

The output loss due to search frictions can be defined as,

$$X \equiv (1 - u) (1 - E_x Y) + u (1 - B) + vK. \quad (4)$$

The output loss is made up of the three components, each of them reflected by a term in equation (4). The loss due to: (i) mismatch, (ii) unemployment, and (iii) the cost of vacancy creation and recruitment. The loss due to mismatch is equal to the employment rate $1 - u$ times the difference between productivity in the optimal assignment, $Y(0) = 1$, and the expected productivity in the actual assignment, $E_x Y$. The loss due to unemployment is equal to the unemployment rate u times the difference between the productivity in the optimal assignment and the value of non-market time $1 - B$. The cost of vacancies

is equal to the vacancy rate v times the cost of a vacancy K .

Two components of this output loss are hard to measure, namely the cost of vacancies

vK and the productivity loss due to mismatch $E_x Y$. However, under free entry, all

profits are spent on vacancy creation so we can substitute equation (??) in. After some

rearrangement, we obtain

$$X = u(1 - B) + (1 - u)(1 - E_x W) = 1 - \rho V^U. \quad (5)$$

The simple relation $X = 1 - \rho V^U$ can be understood easily. Without search frictions,

workers would be costlessly assigned to their optimal assignment where they earn a wage

equal to one and there would be no vacancy cost. Equations (4) and (5) allow us to

estimate the output loss due to search frictions and decompose this loss into its three

components.

Wage formation

The definition of \bar{x} as the upper bound of the mismatch indicator implies that for this level of x , wages are equal to output

$$\widehat{W}(\bar{x}) = \widehat{Y}(\bar{x}) = 1 - \frac{1}{2}\gamma\bar{x}^2. \quad (6)$$

At the marginal job type, all the surplus should go to the worker. If not, the firms would expand their matching set. On the worker side, the definition of \bar{x} being the upper bound of x implies that an unemployed worker is indifferent between accepting this job or staying unemployed. Hence, $V^E = V^U$. Substituting equation (3) in this condition yields

$$\widehat{W}(\bar{x}) = [u + \psi(1 - u)]B + (1 - [u + \psi(1 - u)])E_x W. \quad (7)$$

Since $E_x W > B$, $\widehat{W}(\bar{x}) \geq B$: the lowest wage is greater or equal to the value of leisure.

When $\psi < 1$, a job seeker reduces his chances of finding an even better job by accepting

a job. The excess of the marginal wage offer relative to the value of leisure compensates for this loss in the option value of finding a better job. Only when on- and off-the-job search are equally efficient, $\psi = 1$, equation (7) simplifies to $\widehat{W}(\bar{x}) = B$.

Next, consider the wage for better matches, $0 \leq x < \bar{x}$. We have two cases, one where firms can commit on paying hiring premiums and one where firms cannot; for a full derivation, we refer to Gautier et al. (2010) and web Appendix C.4.

When firms can commit on future wage payments, the optimal wage policy of the firm maximizes the expected value of a vacancy. Even though firms have all the bargaining power, they pay positive wages in order to (i) stimulate new workers to come and (ii) prevent existing workers from quitting. When firms cannot commit on future wage payments, hiring premiums are non-credible since firms would stop paying them as soon as

the worker has accepted the job. Hence, firms only pay no quit premia.

Figure 1 depicts $\widehat{Y}(x)$ and $\widehat{W}(x)$ for both cases with and without commitment. We

use the benchmark values for B, ψ, κ and γ which will be motivated in Section 3.4 below.

Contrary to $\widehat{Y}(x)$, $\widehat{W}(x)$ is non-differentiable at $x = 0$. This is due to the hiring and no-

quit premiums that firms pay. Since the density of employment is highest for low values

of $|x|$, the elasticity of labor supply is high for these types of jobs. A slight variation in

wages has large effects both on the probability that workers accept an outside job offer

and on the number of workers who are prepared to accept the wage offer (the latter being

relevant in the case with commitment only). Hence, firms will bid up wages aggressively

for those types of jobs. Figure 1 shows that the wage in the optimal assignment is higher

when firms can commit than when they cannot, since the ability to commit increases

competition between firms for workers. Figure 1 also reveals that for $x = 0$ the slope of the wage function is smaller (in absolute value) for the case with commitment than without commitment.

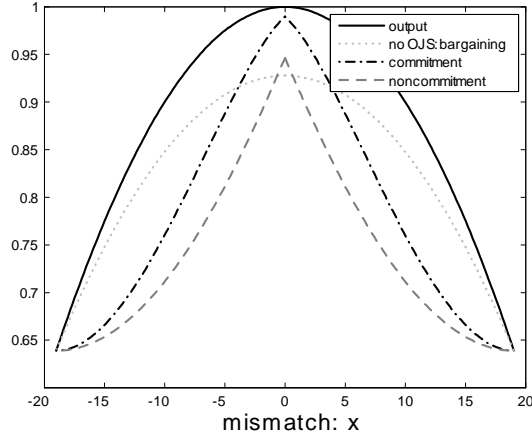


Figure 1: Productivity $\hat{Y}(z)$ and the shape of $\hat{W}(z)$ in different regimes, $B = 0.4, \psi = 0.54, \gamma = 1.8, \kappa$ is chosen such that $u = 5\%$

The dashed line is the wage function that would apply under Nash bargaining without OJS which underlies the analysis in Gautier and Teulings (2006). This wage function does not feature the non-differentiability at $x = 0$. It is just a simple parabola. In this paper

we extend the methodology applied in Gautier and Teulings (2006) for estimating search frictions to the more complex shape of the wage function for the case with OJS. Note that without OJS, we have to assume Nash bargaining, since wage setting by the firm, as is assumed here, would lead to the Diamond paradox of all wage offers being equal to the value of leisure and the full surplus going to the firm. With OJS, competition for workers between firms provides workers with market power (especially for x close to 0) even when they have no bargaining power at all.

The distribution of x among employed workers

The distribution of x among employed workers, $\widehat{G}(x)$, is given by,

$$\widehat{G}(x) = 1 - \frac{\bar{x} - x}{(1 + \psi\kappa x)\bar{x}}. \quad (8)$$

See web Appendix C.1 for a derivation. Figure 2 depicts this distribution function and

the density function that goes with it. The vertical line gives the upper support, \bar{x} . The main message from Figure 2 is that the distribution of x has a large probability mass close to zero (the optimal assignment) and a long right tail of bad matches. The median value of κx is equal to $(1 - u)/(1 + u) < 1$, which is far smaller than the upper support $\kappa \bar{x} = (1 - u)/u$ for reasonable values of u . In fact, 80% of the workers has $x < 5$. The reason for this pattern is that workers who are matched badly quit their jobs fast so their density is low. The reverse holds for good matches, so their density is high. The skewness of the distribution of the mismatch parameter has profound consequences for the difference in wage dispersion between the commitment and no commitment cases.

Equilibrium

The equilibrium can be summarized by three relations as a function of the model's

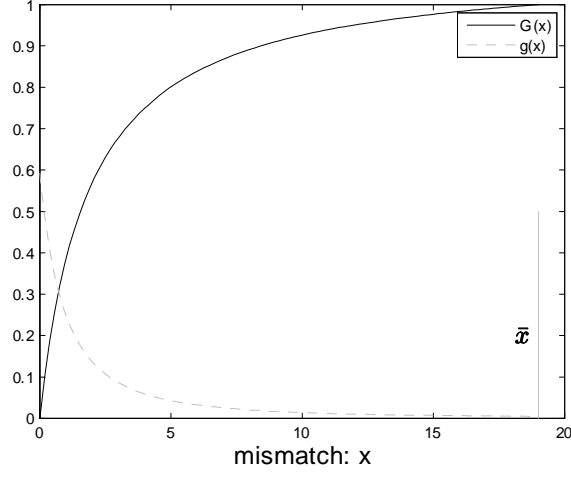


Figure 2: The distribution (solid) and density function (dashed) of x conditional on employment, parameters same as in Figure 1

four parameters B, ψ, κ and γ

$$\widehat{W}(0) - E_x W = \widetilde{W}(B, \psi, \gamma, \kappa), \quad (9)$$

$$\widehat{Y}(0) - E_x Y = \widetilde{Y}(B, \psi, \gamma, \kappa) = \frac{1}{2}\gamma \text{Var}[x],$$

$$u = u(B, \psi, \gamma, \kappa).$$

For all three relations there are two versions, one for the case with commitment and

one for the case without. The analytical expressions for these functions are presented

in Web appendices, C.5, C.1 and C.4. $\widehat{W}(0) - E_x W$, the *max-mean* wage differential,

can be interpreted as a measure of wage dispersion or the wage loss due to mismatch.

For the relevant range of κ , the partial derivatives of $\widetilde{W}(B, \psi, \gamma, \kappa)$, $\widetilde{Y}(B, \psi, \gamma, \kappa)$ and

$u(B, \psi, \gamma, \kappa)$ with respect to κ are negative. Other things equal, an increase in search

frictions (a lower value of κ) leads to more wage dispersion, more output loss due to

mismatch, and more unemployment. These relations lay the foundations for our empirical

inference.

The *max-min* wage differential is larger in the case with commitment because the

maximum is higher, see Figure 1, other statistics of wage dispersion (based on the lowest

wage) may lead to the opposite conclusion, since most of the probability mass of em-

ployment is close to the optimal assignment. In that region, the wage function $W(x)$ is

steeper in the case without commitment, leading to larger wage differentials.⁹

Constrained efficiency

Gautier, Teulings, and Van Vuuren (2010) show that for the special case, $\psi = 1$ (on- and off-the-job search equally efficient) and a quadratic job search technology, $\alpha = 1$, commitment yields constrained efficiency. By the zero profit condition, firms create vacancies till the point that the net present value of expected quasi rents $(1 - u)(E_x Y - E_x W)$ is equal to the cost of a vacancy. Since non-commitment generates higher quasi rents for firms than commitment, non-commitment leads to excess vacancy creation. Cai et.al. (2014) show numerically that the efficiency result for the case of commitment extends to lower values of ψ , $0.25 < \psi < 1$.¹⁰

⁹This is also the reason that the min-mean wage differential used in Hornstein et.al. (2010) is more sensitive to small variations in the models parameters and to measurement error in the minimum. Since our approach uses the max-mean wage differential, it is less sensitive to this problem.

¹⁰For $\alpha < 1$, the standard congestion externalities apply and in that case the commitment case does

3 Estimation

3.1 Measuring mismatch

Our basic strategy for empirical inference is the same as in Gautier and Teulings (2006).

They estimate the output loss due to search frictions for a model without OJS. This

subsection first summarizes their results and then we extend them to the case with OJS.

Gautier, Teulings and Van Vuuren (2005) show that the Taylor approximations of the

search equilibrium in a hierarchical model without OJS correspond one-to-one to the

equilibrium in a circle model. We apply the same analogy here. In this analogy, it makes

sense to talk about a worker's type s as an index of her *skill* and about the job's type c as

an index of its *complexity*. The idea is to establish empirical counterparts for these skill and

complexity indices and then calculate the implied mismatch indicator as the difference

not generate the socially efficient outcome.

between both indicators: $x \equiv s - c$. Skilled workers are assumed to have an absolute advantage in any job type and have a comparative advantage in complex jobs. Hence, in a Walrasian equilibrium, better-skilled workers earn higher wages. Comparative advantage as defined here requires log supermodularity in the production function: better skilled workers are relatively more productive in more complex jobs. Therefore, skilled workers sort into complex jobs and hence wages are an increasing function of job complexity. Those features do not necessarily carry over to a world with search frictions (see Shimer and Smith, 2000), but under a log supermodular production function they do in expectation.¹¹

We use these positive correlations between the worker skill and job-complexity indices on the one hand, and wages, on the other hand to construct indices of worker's skill

¹¹See Gautier and Teulings (2006) footnote 7 for a more detailed analysis of this issue. See also Eeckhout and Kircher (2011).

and job's complexity. For this purpose, we run two regressions: one of demeaned log wages on worker characteristics, like gender, race, years of education and experience, and one on job characteristics, like occupation and industry dummies. Details are in Web appendix C.7. The estimated parameter vector can then be used to construct indices for the observed worker and job characteristics \hat{s} and \hat{c} . Both indices \hat{s} and \hat{c} have zero mean (since wages are demeaned) and are uncorrelated to their unobserved components ε_s and ε_c , respectively. Thus, the skill measure is the predicted log wage conditional on standard worker characteristics and the job complexity level is the predicted log wage conditional on job characteristics. Having estimated both indices, the proxy for x is constructed as $\hat{x} \equiv \hat{s} - \hat{c}$. This way of constructing the skill index \hat{s} implies that the choice of dimension of s is such that the Mincerian rate of return on the skill index is

equal to one: $dE_x[\ln W]/ds = dE_x[\ln W]/d\hat{s} = 1$. Moreover, this specification implies

that $\ln W$ is linear in \hat{s} and \hat{c} . A similar implication holds for the complexity index \hat{c} .

These characteristics are just convenient normalizations of the units of measurements that

imply no loss of generality.

Gautier and Teulings (2006) run wage regressions for six countries where they enter

both \hat{s} and \hat{c} simultaneously, joint with their second order terms

$$\ln W = \bar{\omega}_0 + \omega_s \hat{s} + \omega_c \hat{c} + \omega_{ss} \hat{s}^2 + \omega_{sc} \hat{s} \hat{c} + \omega_{cc} \hat{c}^2 + \varepsilon.$$

Since neither $E[\hat{s}^2]$, nor $E[\hat{s}\hat{c}]$, nor $E[\hat{c}^2]$ are equal to zero, an intercept is added to the

regression. This intercept will play a crucial role. For all six countries $\beta_{ss} < 0, \beta_{sc} > 0$,

and $\beta_{cc} < 0$, and roughly $\beta_{sc} = -2\beta_{ss} = -2\beta_{cc}$. This finding is consistent with the idea

that the final three terms measure the effect of the mismatch indicator $\hat{x}^2 = (\hat{s} - \hat{c})^2$ on

wages. Below we reiterate their result for the United States, which are taken from the

March supplements of the CPS 1989-1992 (t-values in brackets).

$$\ln W = 0.013 + \underset{(8.9)}{0.61\hat{s}} + \underset{(182.4)}{0.66\hat{c}} - \underset{(21.2)}{0.17\hat{s}^2} - \underset{(21.6)}{0.17\hat{c}^2} + \underset{(36.6)}{0.43\hat{s}\hat{c}},$$

$$\ln W = 0.024 + \underset{(14.7)}{0.61\hat{s}} + \underset{(182.2)}{0.66\hat{c}} - \underset{(35.1)}{0.20\hat{x}^2}, \quad (10)$$

$$\text{Var}[\hat{x}] = 0.120. \quad (11)$$

The coefficients on \hat{s} and \hat{c} are between zero and one and highly significant. Since at

the optimal assignment there is a one-to-one correspondence between s and c , we cannot

conclude much from the first order terms. The one could be a proxy for the measurement

error in the other and the other way around. The second-order terms enter also highly

significantly, with the expected signs. In the second regression we impose the restriction

$\beta_{sc} = -2\beta_{ss} = -2\beta_{cc}$. Although a formal F-test rejects them due to the large number of

observations, these restrictions hold almost perfectly. This applies to all six countries.

Gautier and Teulings (2006) provide two arguments why the second order terms are likely to capture the effect of search frictions, which we reiterate here shortly. First, when observed and unobserved worker and job characteristics are distributed jointly normal, it is impossible for second-order terms to be a proxy for the unobserved component of a first-order term, because the correlation of a second-order term in \hat{s} and/or \hat{c} with the unobserved skill index is a third moment and third moments of a normal distribution are equal to zero. Second, the interpretation of these coefficients as capturing the concavity of the wage function implies sign restrictions, which are met for all three coefficients for all six countries. We add one new argument here.¹² If the significance of the second-order terms is indeed driven by the concavity of the wage function in the mismatch indicator, then their

¹²We thank Jean Marc Robin for the idea of this test.

sign would depend on \hat{s} and \hat{c} capturing worker and job characteristics respectively. To the contrary, if both vectors were composed out of mixtures of job and worker characteristics (e.g. experience and occupation dummies in \hat{s} and education and industry dummies in \hat{c}) then the concavity result should not come out. Equation (12) demonstrates this by putting education and occupation in \hat{s} , while equation (13) demonstrates it by putting education and industry in \hat{s} (and the remaining variables in \hat{c}). In both cases, the concavity result disappears

$$\ln W = 0.00 + 0.52\hat{s} + 0.65\hat{c} - 0.01\hat{s}^2 - 0.04\hat{c}^2 + 0.09\hat{s}\hat{c} \quad (12)$$

(0.0) (132.6) (175.2) (0.9) (4.1) (6.2)

$$\ln W = 0.00 + 0.32\hat{s} + 0.81\hat{c} - 0.01\hat{s}^2 + 0.01\hat{c}^2 + 0.05\hat{s}\hat{c} \quad (13)$$

(0.0) (79.9) (233.2) (4.4) (1.1) (3.5)

Also note that the constant moves towards zero in that case. Hence, the concavity result in (10) is not a statistical artifact.

The quadratic terms in equation (10) correspond nicely to the model without OJS, where the wage function $W(x)$ follows a smooth parabola, see Figure 1. Then, the only remaining question is to what extent the coefficient on \hat{x}^2 is affected by measurement error. Here, we want to apply this methodology to a model with OJS. But then we have to find a method for fitting a wage function that is not a simple parabola, but a more complicated function that is non-differentiable at $x = 0$. This question will be addressed below.

3.2 Capturing the shape of $W(x)$

The first step in finding a tractable approach to estimating the function $\ln W(x)$ is to consider a simple Taylor expansion around the optimal assignment, $x = 0$. In section 4 when we calibrate the model, we will use the exact expressions and test how well the

approximations below perform. Start from the circle model. Then

$$\ln W(x) = \omega_0 - \omega_2 |x| + O(x^2), \quad (14)$$

where we demean the data on $\ln W$, such that $E_x \ln W = 0$. The wage curves in Figure

1 imply that $\omega_2 > 0$. If there were no search frictions, then $x = 0$ at all job types and

therefore $\omega_0 = 0$. By construction, ω_0 is approximately equal to our measure of wage

dispersion, the *max-mean* wage differential.

$$\omega_0 = \omega_2 E[|x|] = \ln W(0) - E_x \ln W \gtrsim \ln W(0) - \ln E_x W \gtrsim W(0) - E_x W. \quad (15)$$

The first equality follows from taking expectations at the left- and right-hand sides of (14)

and using $E_x \ln W = 0$ (by demeaning). The second equality follows from evaluating (14)

at $x = 0$. The next approximate equality is due to Jensen's inequality, $\ln E_x W \gtrsim E_x \ln W$.

For the third step, note that for small search frictions –and accordingly, small wage

differentials— $E_x W \lesssim W(0) \lesssim 1$, the approximation $\ln W \lesssim W - 1$ applies. Finally, since

$W(0) \simeq 1$, $W(0) - E_x W \simeq [W(0) - E_x W] / W(0)$. Hence, ω_0 is a convenient statistic for

the relative wage loss due to search frictions.

In practice we observe x with a fair amount of measurement error. What are the implications of this? Let $\varepsilon_x \equiv \varepsilon_c - \varepsilon_s$ be the measurement error in the observed signal \hat{x} ,

$$\hat{x} = x + \varepsilon_x, \tag{16}$$

with $\text{Cov}[x, \varepsilon_x] = 0$.¹³ Hence, $\text{Var}[\hat{x}] = \text{Var}[x] + \sigma_\varepsilon^2$, where $\sigma_\varepsilon^2 \equiv \text{Var}[\varepsilon_x]$. Measurement

error is particularly relevant when estimating the effect of mismatch, since the observed

¹³By construction, $\text{Cov}[\varepsilon_s, \hat{s}] = \text{Cov}[\varepsilon_c, \hat{c}] = 0$. Hence, ε_s and ε_c measure unobserved heterogeneity in s and c . This does not apply to the mismatch indicator \hat{x} , where $\text{Cov}[\varepsilon_x, \hat{x}] > 0$. This can be seen most easily by considering the limiting case of zero search frictions (the Walrasian equilibrium), where $\text{Var}[x] = 0$ and $\text{Var}[\hat{x}] = \text{Var}[\varepsilon_x] \geq 0$, since $s = c$ and hence $\varepsilon_x = \hat{x}$. For small search frictions, $\text{Var}[x] \ll \text{Var}[\varepsilon_x]$, $\text{Cov}[\varepsilon_x, x] \cong 0$. Hence, ε_x can be interpreted as pure classical measurement error in the observed mismatch indicator \hat{x} . The signal-to-noise ratio that we find empirically supports this interpretation, see Section 4.

mismatch \hat{x} can either be due to true mismatch or to measurement error. In a perfect Walrasian world, there is no mismatch, since $s = c$ for each job. Hence, $x = 0$. A careless researcher would 'observe' mismatch since the observed skill and complexity indexes are not equal, $\hat{s} \neq \hat{c}$, due to unobserved heterogeneity in s and c . Hence, the 'observed' mismatch \hat{x} is equal to the measurement error: $\hat{x} = \hat{s} - \hat{c} = \varepsilon_x$ since $x = 0$ for each job. Failing to correct for the impact of measurement error will therefore overestimate the importance of mismatch.

Define the signal-to-noise ratio $R \equiv \text{Var}[x] / \text{Var}[\hat{x}]$. If the approximation of $\ln W(x)$ in equation (14) would be exact, estimating this equation with OLS (replacing $|x|$ by $|\hat{x}|$) would yield a downwardly-biased estimate of ω_2 for two reasons. First, attenuation bias due to measurement error biases the coefficient on the explanatory variables towards zero;

and second, there is the strong convexity at zero. Due to this convexity,

$$\mathbf{E}[|x| \mid \hat{x}] \geq |\hat{x}|.$$

The closer \hat{x} is to zero, the stronger this inequality. This is documented in Figure 3, where

we present three functions, $|\hat{x}|$, $\mathbf{E}[|x| \mid \hat{x}]$, and the least-squares estimation of $\mathbf{E}[|x| \mid \hat{x}] =$

$\beta_0 + \beta_2 \hat{x}^2 + \varepsilon$, for the case that both the true value x and measurement error ε_x are

normally distributed –both with variance equal to unity. The least-squares approximation

of $\mathbf{E}[|x| \mid \hat{x}]$ turns out to be extremely precise for the relevant range between plus- and minus

two standard deviations of $\hat{x} \in [-2, 2]$. This justifies the idea of approximating equation

(14) by a regression model of w with a quadratic term \hat{x}^2 :

$$\ln W = \bar{\omega}_0 - \bar{\omega}_2 \hat{x}^2 + \varepsilon, \tag{17}$$

where ε is a zero mean error term. This is a surprising result: while the wage function

$W(x)$ with OJS is entirely different from the wage function without OJS, a second order

polynomial is again an accurate approximation of the relation between log wages and the

observed proxy for mismatch, \hat{x} in the presence of measurement error.

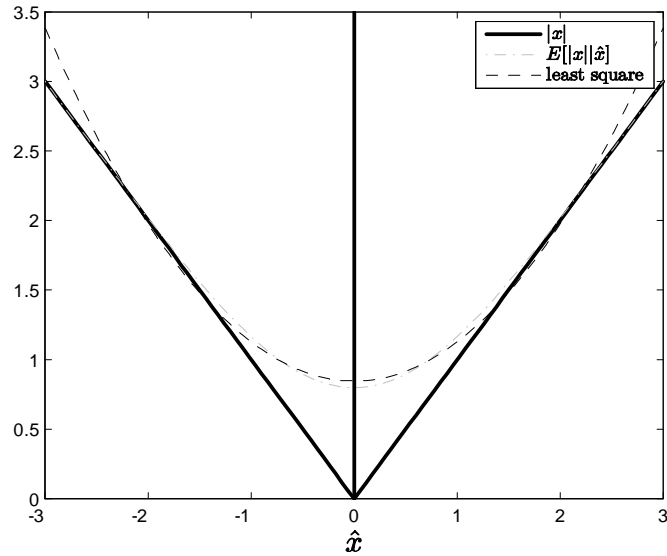


Figure 3: Smoothing of an absolute value function by random mixing for $\sigma_x^2 = \sigma_\varepsilon^2 = 1$:

$|x|$ (black thin), $E[|x| | \hat{x}]$ (blue dotted), least- squares estimate (red solid)

The following proposition relates the least-squares estimate $\bar{\omega}_0$ to the underlying co-

efficient ω_0 .

Proposition 1 *Suppose (i) that the true model is given by*

$$\ln W = \omega_0 - \omega_2 |x|,$$

where both $\ln W$ and x are normalized to have a zero mean, (ii) that we observe only

$\hat{x} = x + \varepsilon_x$ where both x and ε_x are distributed normally with $\text{Var}[x] / \text{Var}[\hat{x}] \equiv R$, (iii)

that we estimate equation (17) by OLS. Then,

$$\text{plim} \bar{\omega}_0 = \frac{1}{2} R \omega_0 = \frac{1}{2} R \omega_2 E|x|.$$

Proof: See Appendix A.3.

Hence, when there is no measurement error in the observed signal \hat{x} ($R = 1$), the estimated intercept $\bar{\omega}_0$ is equal to half the true intercept ω_0 . This underestimation by a factor two is due to the imperfect approximation of the absolute value function by a parabola. When on top of this imperfection in the functional form, there is also measure-

ment error in the signal \hat{x} , the underestimation becomes more severe. However, as Figure 3 shows, a parabola provides a very good description for $E[|x|]$ when x is convoluted with measurement error, in particular when $R < 1/2$. Hence, using a parabola is an efficient way of estimating ω_0 . Given the measurement error in the data, one cannot do much better by using alternative estimation methods. The estimate of $\bar{\omega}_0$ is proportional to the signal-to-noise ratio R . Under the assumption of joint normality of \hat{x} and ε_x , Proposition 2 and equation (15) imply that the intercept $\bar{\omega}_0$ underestimates the true magnitude of wage dispersion by a factor $2/R$:

$$\frac{W(0) - E_x W}{W(0)} \cong \frac{2}{R} \bar{\omega}_0. \quad (18)$$

Equation (18) provides a very convenient relation. Twice the intercept of a simple OLS regression, $2\bar{\omega}_0$, provides a robust estimate for the magnitude of the wage loss due to search

frictions. The estimation results in equation (10) imply $\bar{\omega}_0 = 0.024$. Hence, the relative wage loss due to mismatch $[W(0) - E_x W]/W(0)$ is at least 4.8%. This is the limiting case of $R = 1$, when there is no measurement error. In the presence of measurement error, the wage loss due to mismatch is larger.

Since $\text{Var}[x] = R\text{Var}[\hat{x}]$, the productivity loss due to mismatch follows from

$$Y(0) - E_x Y = \frac{1}{2}\gamma E[x^2] = \frac{1}{2}\gamma R\text{Var}[\hat{x}].$$

Conditional on γ , the variance of the observed mismatch indicator $\text{Var}[\hat{x}]$ overestimates the productivity loss due to mismatch by a factor R , exactly the reverse of the underestimation of the expected wage dispersion. The latter is due to the fact that part of the variance of a noisy mismatch indicator does not reflect true mismatch, but just noise. The elimination

of R from these two expressions yields

$$[W(0) - E_x W][Y(0) - E_x Y] = \widetilde{W}(B, \psi, \gamma, \kappa) \widetilde{Y}(B, \psi, \gamma, \kappa) \cong \gamma \bar{\omega}_0 \text{Var}[\widehat{x}]. \quad (19)$$

Equation (19) is a key equation. It establishes a relation between the four parameters of

the model, B, ψ, κ and γ on the left hand side and the parameter γ and the estimated

statistics $\bar{\omega}_0$ and $\text{Var}[\widehat{x}]$ on the right hand side. For given γ , the product of wage and

productivity dispersion is not affected by measurement error. More measurement error in

x , increases $\text{Var}[\widehat{x}]$ by the same factor as it decreases $\bar{\omega}_0$.

3.3 Measuring γ and its role in identification

Teulings (2005) shows that there exists a one-to-one correspondence between the Katz

and Murphy (1992) elasticity of complementarity between low and high skilled workers,

$\eta_{\text{low-high}}$, and the parameter γ .¹⁴ Katz and Murphy estimate this elasticity to be 1.4 for the period 1963-87: a 1% increase in the ratio between high- and low-skilled workers yields a 1.4% fall in the relative wages of high-skilled workers. Suppose that the data are generated by a continuous type model like ours, but that the researcher arbitrarily divides the workforce into two groups, high and low skilled, where all workers below a certain threshold value for s are assigned to the low skilled group and all workers above that threshold are assigned to the high skilled group. When this researcher then tries to estimate $\eta_{\text{low-high}}$, he will obtain the following result:

$$\gamma = \frac{1}{\text{Var} [\ln W] \eta_{\text{low-high}}} \cong \frac{1}{0.40 \times \eta_{\text{low-high}}}.$$

¹⁴In the web appendix C.6, we show that the units of s ($dE[\ln W]/ds = 1$) are the same as in Teulings (2005) so we can transfer his findings to this paper.

Katz and Murphy (1992)'s benchmark value for $\eta_{\text{low-high}}$ of 1.4 yields (using the empirical value for $\text{Var}[\ln W]$ of 0.40) $\gamma = 1.8$. Their discussion on pages 71-72 suggests that for the period 1975-87, choosing $\eta_{\text{low-high}} = 4$, performs better than 1.4. The alternative value for $\eta_{\text{low-high}}$ of 4 yields $\gamma = 0.6$. Better substitutability of worker types reduces the output loss due to mismatch.

Figure 4 illustrates the situation for both the commitment and non-commitment cases, using the benchmark values for B , ψ , and γ from the next section. The upward-sloped curve (red) consists of combinations of wage loss $\widetilde{W}(B, \psi, \gamma, \kappa)$ and productivity loss $\widetilde{Y}(B, \psi, \gamma, \kappa)$ from equation (9). Each point on the curve corresponds to different values of κ . When there are more search frictions (low κ), there is both more wage dispersion and more mismatch. The downward-sloped curve (blue) reflects equation (19), using

the values of $\bar{\omega}_0$, and $\text{Var}[\hat{x}]$ discussed above. This yields a hyperbolic relation between

$\widetilde{W}(B, \psi, \gamma, \kappa)$ and $\widetilde{Y}(B, \psi, \gamma, \kappa)$. Each point on this curve corresponds to a different value

of the signal-to-noise ratio R . In the North-West, $R = 1$ and $\widetilde{W}(B, \psi, \gamma, \kappa) = 2\bar{\omega}_0 = 4.8\%$.

Note that R is not a completely free parameter because the model restricts the curves

to intersect at values for $R \leq 1$. The lower R , the more $\text{Var}[\hat{x}]$ overestimates $\text{Var}[x]$ and

hence $\widetilde{Y}(B, \psi, \gamma, \kappa)$, but the more $\widetilde{W}(B, \psi, \gamma, \kappa)$ is underestimated. The intersection of

both curves determines κ and R .

Wage differentials are generally larger than productivity differentials, which is due to

the fact that the derivative of the production function is zero in the optimum, $Y'(0) = 0$,

while the wage function is non-differentiable at that point. Since $x = 0$ is the point

with the highest density, this point matters a lot for the relative size of productivity

and wage differentials.¹⁵ The ultimate test of the model is to see whether the implied value of unemployment is realistic. Figure 5 plots the relation between $\widetilde{W}(B, \psi, \gamma, \kappa)$ and $u(B, \psi, \gamma, \kappa)$. The two points on the curves in Figures 4 and 5 denote the exact solutions, which will be discussed in Section 4. A higher value of γ shifts the locus of equation (19) to the North-East in Figure 4. For a given amount of wage dispersion, a higher γ implies more output loss due to mismatch ($\frac{1}{2}\gamma\text{Var}[\widehat{x}]$). At the intersection with the upward sloped locus of equation (9), both $\widetilde{W}(B, \psi, \gamma, \kappa)$ and $\widetilde{Y}(B, \psi, \gamma, \kappa)$ are higher. Figure 5 shows that the higher value of $\widetilde{W}(B, \psi, \gamma, \kappa)$ that corresponds to $\gamma = 1.8$, implies a value of the unemployment rate around 10% (the exact solutions imply a lower u).

¹⁵Note that for the model without OJS in Gautier and Teulings (2006), productivity differentials are always larger than wage differentials.

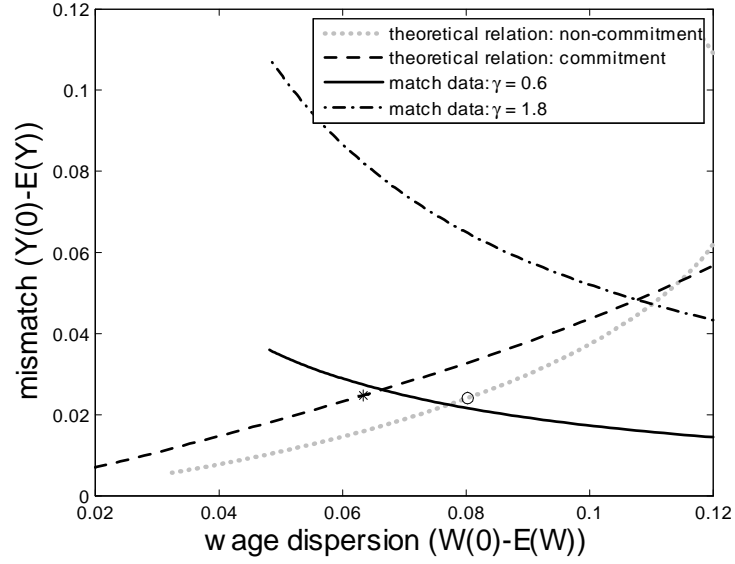


Figure 4: Identification, the role of γ ($B = 0.4$, $\psi = 0.54$)

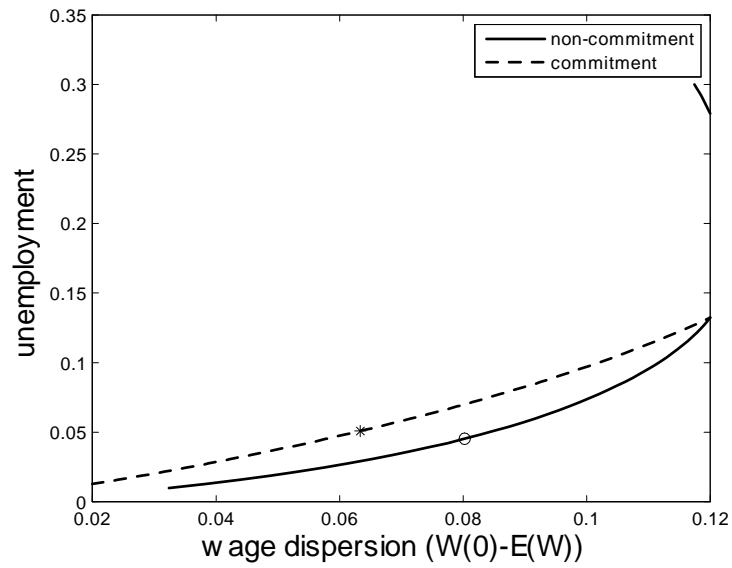


Figure 5: Unemployment ($B = 0.4$, $\psi = 0.54$)

3.4 The values of B and ψ

Hall and Milgrom derive a value for B based on UI benefits of 0.25 and an estimated

Frisch elasticity of labor supply of 1. This implies a value of $B = 0.71$.¹⁶ In our setting,

where we are interested in the mean or median worker, such a high value would have the

unpleasant implication that if nobody would work in the US, per capita output would

only be 27% lower than in case everybody would be at her optimal job. Therefore, we

think it is more reasonable to set $B = 0.4$ following Shimer (2005). As a robustness check,

we also calibrate our model for $B = 0.6$.

The value of ψ is identified from the relation between the ratio of the (average)

¹⁶Hagedorn and Manovskii (2008) and Hall (2009) seek to explain the cyclical behavior of unemployment, so they use larger values for B . For these studies, the value of non-market time of the *marginal worker* is relevant whereas here we are interested in the value of non-market time for the *average worker*, which justifies a lower value of B .

employment-to-employment hazard rate, f_{ee} , and the employment-to-unemployment hazard rate, f_{eu} , see Appendix A.2

$$\frac{f_{ee}}{f_{eu}} = \frac{1+q}{q} \ln(1+q) - 1, \quad (20)$$

$$q \equiv \psi \frac{1 - u(B, \psi, \gamma, \kappa)}{u(B, \psi, \gamma, \kappa)}.$$

Since q is an increasing function of f_{ee}/f_{eu} , its inverse exists. Denote this inverse by

$\Psi(f_{ee}/f_{eu})$. Hence, $\psi = \frac{u(B, \psi, \gamma, \kappa)}{1 - u(B, \psi, \gamma, \kappa)} \Psi(f_{ee}/f_{eu})$. Identification of the model proceeds along

the same lines as in Figure 4, but now taking B and f_{ee}/f_{eu} as given instead of B and ψ .

Monthly transition rates from 1967-2010 similar to Shimer (2007) imply that $f_{eu} = 2.13\%$ per month for the mean worker. The value of f_{ee} is 2.7% according to Fallick and Fleischman (2004), 2.9% according to Nagypal, and 3.3% according to Moscarini and

Vella (2008), applying a correction for missing records in the CPS.¹⁷ Hence, for the mean

¹⁷Nagypal's values come from the SIPP and the CPS. She argues that those estimates are down-

worker, $\frac{2.7}{2.13}(= 1.27) \leq f_{ee}/f_{eu} \leq \frac{3.3}{2.13}(= 1.55)$.¹⁸

If we consider the median worker, a lower value for f_{eu} applies. This can be seen as follows. According to the BLS statistics, median tenure is 4.6 years.¹⁹ In the absence of duration dependence and ignoring the flow out of the labor force, the total hazard out of the current job, $f_{ee} + f_{eu}$, is 1.3%.²⁰ The transition rate f_{eu} (the equivalent of δ) is assumed to be constant in our model, while the unconditional transition rate, f_{ee} , exhibits negative duration dependence due to heterogeneity in the match quality x : high

quality matches survive. Negative duration dependence implies that the hazard rate for wardly biased because when workers change jobs it is not uncommon for them to experience a short unemployment spell. In the data, this yields an employment-unemployment transition followed by an unemployment-employment transition. This bias might be larger than the time aggregation bias in the unemployment outflow rate.

¹⁸We thank Bart Hobijn for sharing his data.

¹⁹www.bls.gov/news.release/pdf/tenure.pdf

²⁰Ignoring flows out of the labor force, the total hazard out of employment can be solved from $1 - \exp[-55(f_{ee} + f_{eu})] = 0.5$.

low-tenure workers is above 1.3% and the rate for high-tenure workers is below 1.3%.

Since f_{eu} is constant, it must be smaller than 1.3%, much lower than the value reported

by Shimer. This implies that the assumption of the absence of duration dependence of f_{eu}

is rejected by the data. Apparently, a small group of weakly attached workers frequently

transits between unemployment and employment. In order to capture this feature of

reality, other mechanisms must be introduced (such as learning, see Moscarini (2005);

or random growth, see Buhai and Teulings (2014)). This falls outside the scope of this

paper. Hence, our model is unable to explain this feature of reality. We set $f_{ee}/f_{eu} = 1.75$

in our preferred calibration, but we check the robustness of our results for higher values

of f_{ee}/f_{eu} in Appendix B.

4 Calibration

The methodology applied in the previous section requires two approximations. First, we approximate $\ln W(x)$ by a first order Taylor expansion in Section 3.2, using the absolute value transformation: $-\omega_2 |x|$. Second, we approximate the distribution of x by the normal distribution, see Proposition 2, while its actual distribution is far from normal, see Figure 2. In this section, we use the exact expressions to calibrate our model. The calibration proceeds as follows,

1. We set $B = 0.4$ and $\gamma = 1.8$ (our preferred values, see the discussion in the previous section).
2. Take starting values for ψ and κ .
3. $\text{Var}[x]$ can be calculated directly, see (31) in Web Appendix C.5. The variance of

the measurement error distribution, σ_ε^2 , can be calculated from

$$\sigma_\varepsilon^2 = \text{Var}[\hat{x}] - \text{Var}[x],$$

using the empirical value $\text{Var}[\hat{x}] = 0.12$. Note that consistency requires $\text{Var}[\hat{x}] \geq$

$\text{Var}[x]$, which is an additional test for the model (i.e. the intersection of the blue and

red curve in Figure 4 occurs before the end point of the blue curve that corresponds

to $R = 1$ in the north west). This condition turns out to hold in our calibration.

4. The f_{ee}/f_{eu} ratio in the model follows from (20) while the simulated $\bar{\omega}_0$ is obtained

as follows. First draw values from $\hat{G}(x)$, see (8). Next, add measurement error

using the value of σ_ε^2 from step 3 and run regression (17). Compare the simulated

values for f_{ee}/f_{eu} and $\bar{\omega}_0$ to the empirical values, $f_{ee}/f_{eu} = 1.75$ and $\bar{\omega}_0 = 0.0241$.

As long as they do not match, adjust ψ and κ and return to step 2 till convergence

is reached.

This procedure converges fast (also for the alternative calibration in Web Appendix

B).

Table 1 presents the implied values for our key variables.

	$u(\%)$		ψ		$W(0) - E_x W$ x100		$E_x W / W(\bar{x})$		$Y(0) - E_x Y$ x100		$R(\%)$	
commitment	yes	no	yes	no	yes	no	yes	no	yes	no	yes	no
	4.72	4.48	0.58	0.54	6.04	8.01	1.51	1.44	2.32	2.39	21.5	22.2

Table 1: Calibration results for $B = 0.4, fee/feu = 1.75, \gamma = 1.8$

The implied unemployment rate in Table 1 is in the reasonable range for both the commitment (4.72%) and non-commitment case (4.48%). Hence, the data do not allow us to discriminate between commitment and non-commitment. On-the-job search is about half as efficient as off-the-job search. The mean-min ratio predicted by the model is similar to the one reported by Hornstein et al. (2010) if they use the 10th percentile as the lowest

wage. The signal-noise ratio R may appear low at first sight. However, remember that

R would be 0 in the Walrasian case, see the discussion regarding equation (16). Hence,

there is no natural lower bound for R .

Our results also show that the approximations in Section 3.2 overestimated unemploy-

ment and wage dispersion. For the no commitment case, the approximations suggested

that, $W(0) - E_x W = 11\%$ and $u = 10\%$ while for the commitment case, the approxi-

mations would give $W(0) - E_x W = 11\%$ and $u = 11\%$. In the simulations where we use

the exact values we get less wage dispersion ($W(0) - E_x W$ equals for commitment and no

commitment respectively 6% and 8%) and a lower unemployment rate of 4.5%. The high

values in the approximations were mainly due to the fact that there we assumed that x

followed a normal distribution rather than the actual distribution $\hat{G}(x)$ that we use here

(which is far from normal). The linear approximation of the wage function applied in (14)

had little effect.

In Tables 3 and 4 in Appendix B we present the calibration results for 8 other combinations of the parameters B , f_{ee}/f_{eu} , and γ . The calibration with $B = 0.6$, $\gamma = 1.8$ and $f_{ee}/f_{eu} = 1.3$ gives an implausibly large frictional unemployment rate of 11%. In general, for $B = 0.6$, the unemployment rate is a bit higher and for $\gamma = 0.6$, it is a bit lower than for our baseline analysis. ψ can increase to almost 1 if we use $B = 0.6$, $f_{ee}/f_{eu} = 1.75$ and $\gamma = 1.8$. For the no commitment case, unemployment varies between 2.8 and 12.7 for the eight different configurations.

Figure 6 illustrates what happens if we use different parameters in the calibration. For example, if we calibrate the model with a higher value of B , this makes the theoretical

relationship between wage dispersion and mismatch steeper (the yellow curve). With a higher B , it is harder to generate wage dispersion. A given level of wage dispersion is then associated with more frictions and consequently more mismatch. However, at the new intersection, the corresponding unemployment rate and ψ are no longer consistent with the empirical value of f_{ee}/f_{eu} . Therefore, ψ must increase to shift the yellow curve back and this is depicted by the green dotted curve. Figure 7 gives the corresponding values of unemployment.

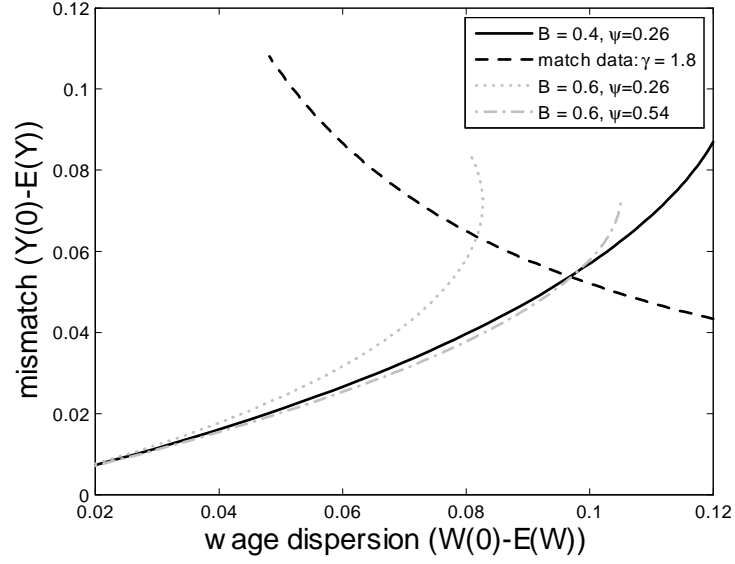


Figure 6: The effect of B on the estimated value of ψ

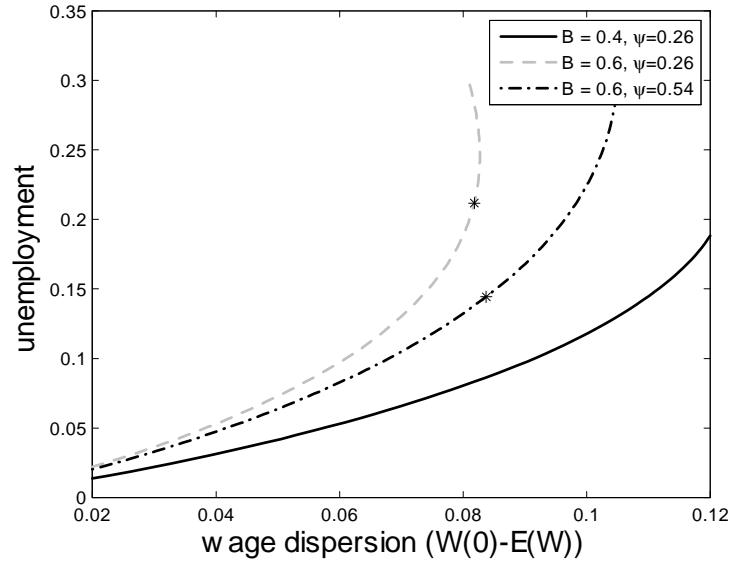


Figure 7: The effect of B on the estimated value of u

4.1 Composition of the output loss and the business-stealing externality

Table 2 shows for our baseline parameters that if firms can commit to wages, the output loss, X , due to search frictions is 10.0% while if firms cannot commit, it is 15.5%. A large share of the output loss in the non-commitment case is due to vacancy creation. This is due to the fact that we assume free entry, or equivalently, an infinite elasticity of vacancy creation with respect to profits per worker. In reality, there will be imperfect competition and part of the profits will be captured by the owners of the firm.

Commitment	yes	no
$u(1 - B)$	2.84	2.69
$(1 - u) [E_x Y - E_x W]$	4.43	10.48
$(1 - u) [Y(0) - E_x Y]$	2.21	2.28
X (%)	9.48	15.45

Table 2: Decomposition of output loss due to frictions for $B = 0.4$, $\gamma = 1.8$ and $f_{ee}/f_{eu} = 1.75$

The large value of $(1 - u)(E_x Y - E_x W)$ for the non-commitment case is socially inefficient because of a business-stealing externality (see Gautier et al. 2010). The idea is that without commitment, when opening a vacancy, individual firms do not internalize the future output loss of the firm from which they poach a worker. Although the transitions of workers to better matches are always efficient, the expected productivity gains are too small to justify, from a social point of view, the entry cost of the marginal firm. Tables 5 and 6 in Appendix B present the estimated output loss for the eight different calibrations and find that X varies between 6.42% (commitment, $B = 0.4$, $\gamma = 0.6$, $f_{ee}/f_{eu} = 1.75$) and 18.78% (non-commitment, $B = 0.6$, $\gamma = 1.8$, $f_{ee}/f_{eu} = 1.3$).

Estimating the business-stealing externality

Table 2 is not suitable to estimate the business-stealing effect because it keeps constant

the outcome variable f_{ee}/f_{eu} , and not the parameter κ . In order to estimate the business-stealing effect, we use for the commitment case the same parameter values, $B = 0.4$, $\gamma = 1.8$, and $\psi = 0.54$ as in the no-commitment case. Under commitment, there is no excessive vacancy creation (vK reduces to 4.67) and this makes the unemployment rate slightly higher (5.05%). The output loss due to mismatch under commitment is almost the same (2.33) as for the original calibration. The total output loss due to search frictions is now 15.45% for no commitment and 10.03% for commitment. The difference of almost 5.5% points is the welfare loss due to the existence of a business-stealing externality that arises if firms cannot commit to wages contingent on x . Note that this estimate is based on the assumption that all excessive rents of the inefficient wage mechanism are spent on vacancy creation. If the rents end up at the firm owners, the losses will be smaller, since

they will derive utility from this income.

5 Conclusion

Due to frictions, only a subset of the contacts between vacancies and workers results in a match, and this creates (i) unemployment, (ii) wage dispersion amongst identical workers and (iii) mismatch. This paper contributes to the literature by measuring these manifestations of search frictions and presenting a model that can jointly explain them (allowing for measurement error). Our methodology yields a very simple and tractable method for estimating wage dispersion due to search frictions using a simple OLS regression on worker and job characteristics. We use the analogy to hedonic pricing models to derive the curvature of the production function from Katz and Murphy's estimate of the elasticity of complementarity between high and low skilled workers. The output loss due to search

frictions only depends on four parameters: the value of non-market time B , the relative efficiency of on-the job search, ψ and a composite parameter that captures everything that affects frictions (the efficiency of matching, the discount rate, the job destruction rate, vacancy creation cost and a parameter measuring the cost of mismatch).

Search frictions generate output losses directly due to the suboptimal allocation of resources, and indirectly, because decentralized wage mechanisms potentially come with distortions. Allowing for two-sided heterogeneity is extremely important because it is the interaction between the search frictions, the type distributions and the production technology that determines how important these frictions are. If workers and firms are identical, then all contacts result in a match. Under two-sided heterogeneity, the production technology matters because it determines how much output is lost due to mismatch.

The more difficult it is to substitute between worker types, the greater this output loss.

By combining information on wage dispersion and the substitutability of worker types we

can learn about the actual amount of frictions and the importance of a precise match.

We then use our model to quantify and decompose this total output loss. Traditionally,

most of the macro labor literature has focussed on unemployment, but our results imply

that mismatch and job creation cost are also important. We find that this total loss is

between 9% and 16%, depending on whether firms can or cannot commit to wages, on the

value of non-market time and on the efficiency of on– relative to off-the-job search. Gau-

tier and Teulings (2006) did not allow for on-the-job search, and therefore substantially

overestimated the output loss due to frictions.

References

- [1] ABOWD, JOHN M., FRANCIS KRAMARZ AND DAVID N. MARGOLIS (1998), High wage workers and high wage firms, *Econometrica*, 67, 251-333.

- [2] ATAKAN, ALP (2006), Assortative matching with explicit search costs, *Econometrica* 74, 667–680.

- [3] BAGGER JESPER AND RASMUS LENTZ, An Empirical Model of Wage Dispersion with Sorting, Mimeo University of Wisconsin-Madison.

- [4] BARTOLUCCI CRISTIAN AND FRANCESCO DEVICIENTI (2012), Better workers move to better firms: A simple test to identify sorting, mimeo Collegio Carlo Alberto.

- [5] BECKER, GARY S. (1973), A theory of marriage: Part I, *Journal of Political Economy* 81(4), 813-46.

- [6] BONTEMPS, CHRISTIAN. JEAN MARC ROBIN AND GERARD J. VAN DEN BERG (2000), Equilibrium search with continuous productivity dispersion: Theory and non-parametric estimation, *International Economic Review*, 41, 305-58.

- [7] BUHAI, I. SEBASTIAN AND COEN N. TEULINGS (2014), Tenure profiles and efficient separation in a stochastic productivity model, *Journal of Business and Economic Statistics*, forthcoming.

- [8] BOUND, JOHN AND ALAN KRUEGER (1991), The extent of measurement error in longitudinal earnings data: Do two wrongs make a right?, *Journal of Labor Economics*, 9, 1-24.

- [9] BURDETT KENNETH AND DALE MORTENSEN (1998), Equilibrium wage differentials and employer size, *International Economic Review*, 39, 257-274.
- [10] CAI XIAOMING , PIETER A. GAUTIER, COEN N. TEULINGS AND MAKOTO WATANABE (2014) Collective versus decentralized wage bargaining and the efficient allocation of resources (2013), *Labour Economics*, 26, 34-42.
- [11] COLES MELVYN (2001), Equilibrium wage dispersion, firm size, and growth, *Review of Economic Dynamics*, 4-1,159-187.
- [12] CORNFELD, OFER (2012), The price of talent, mimeo Tel Aviv University.
- [13] DICKENS, WILLIAM T. AND LAWRENCE F. KATZ (1987), Inter-industry wage differences and theories of wage determination, *NBER* working paper No. 2271.

- [14] EECKHOUT, JAN AND PHILIPP KIRCHER (2011), Identifying sorting, in theory,

Review of Economic Studies 78-3, 872-906.
- [15] EECKHOUT, JAN AND PHILIPP KIRCHER (2010). Sorting and decentralized price

competition, *Econometrica* 78-2, 539-574.
- [16] ELLIOTT, MATTHEW (2010), Inefficiencies in networked markets, mimeo Stanford.
- [17] ELSBY, MIKE, BART HOBIJN AND AYSEGUL SAHIN (2008), Unemployment Dy-

namics in the OECD, NBER Working Paper 14617.
- [18] FALICK, BRUCE AND CHARLES A. FLEISCHMAN (2004), The importance of

employer-to-employer flows in the U.S. labor market, working paper, Federal Re-

serve Bank Board of Governors, Washington DC.

- [19] GAUTIER, PIETER A. AND COEN N. TEULINGS (2006), How large are search frictions? *Journal of the European Economic Association*, 4-6, 1193-1225.
- [20] GAUTIER, PIETER A., COEN N. TEULINGS AND AICO P. VAN VUUREN (2005), Labor market search with two-sided heterogeneity: hierarchical versus circular models, in: H. Bunzel, B.J. Christensen, G.R. Neumann, and J.M. Robin, eds., *Structural Models of Wage and Employment Dynamics, conference volume in honor of Dale Mortensen*, Elsevier, Amsterdam.
- [21] GAUTIER, PIETER A, COEN N. TEULINGS AND AICO P. VAN VUUREN (2010), On-the-Job search mismatch and efficiency, *Review of Economic Studies*, 77-1, 245-272.
- [22] HALL, ROBERT E. (2009), Reconciling cyclical movements in the marginal value of

time and the marginal product of labor, *Journal of Political Economy*, 117, 281-322.

- [23] HALL, ROBERT E. AND PAUL R. MILGROM (2008), The limited influence of unemployment on the wage bargain, *American Economic Review*, 98(4), 1653-1674.

- [24] HAGEDORN MARCUS AND IOURII MANOVSKII (2008), The cyclical behavior of equilibrium unemployment and vacancies Revisited, *American Economic Review*, 98, 1692-1706.

- [25] HAGEDORN MARCUS. TZUO HANN LAW AND IOURRI MANOVSKII (2012), Identifying sorting, mimeo UPenn.

- [26] HORNSTEIN, ANDREAS, PER KRUSELL AND GIOVANNI L. VIOLANTE (2010), Frictional wage dispersion in search models: A quantitative assessment, *American Economic Review*, 101(7), 2873-98.

- [27] JOVANOVIĆ BOYAN (2014), Misallocation and growth, *American Economic Review*, 104, 1149-71.
- [28] KATZ, LAWRENCE AND KEVIN M. MURPHY (1992), Changes in relative wages, 1963-1987: Supply and demand factors, *Quarterly Journal of Economics*, 107, 35-78.
- [29] LISE, JEREMY, COSTAS MEGHIR, AND JEAN MARC ROBIN (2012), Matching, sorting and wages, UCL mimeo.
- [30] LISE, JEREMY, AND JEAN MARC ROBIN (2013), The Macro-dynamics of Sorting between Workers and Firms, UCL mimeo.
- [31] LOPES DE MELO RAFAEL (2008), Sorting in the labor market: theory and measurement, mimeo Chicago.

- [32] MARIMON RAMON AND FABRIZIO ZILIBOTTI (1999), Unemployment vs. mismatch of talents: Reconsidering unemployment benefits, *Economic Journal*, 109, 266-291.
- [33] MOSCARINI, GIUSEPPE (2005), Job matching and the wage distribution, *Econometrica*, 73(2), 481-516.
- [34] MOSCARINI, GIUSEPPE AND FRANCIS VELLA (2008), Occupational mobility and the business cycle, NBER working paper 13819.
- [35] MORTENSEN, DALE T. (2000), Equilibrium unemployment with wage posting: "Burdett-Mortensen meets Pissarides", in B.J. Christensen, P. Jensen, N. Kiefer and D.T. Mortensen, eds., *Panel Data and Structural Labor Market Models*, Amsterdam, Elsevier science.
- [36] MORTENSEN, DALE T. AND CHRIS PISSARIDES (1999), New developments in mod-

els of search in the labor market, *in*: O.C. Ashenfelter and D. Card, *Handbook of Labor Economics* 3b, North-Holland, Amsterdam.

[37] NAGYPAL, EVA. (2005), On the extent of job-to-job transitions, working paper, Northwestern University.

[38] PISSARIDES, CHRIS A. (1994), Search unemployment with on-the-job search, *Review of Economic Studies*, 61, 457-475.

[39] PISSARIDES, CHRIS A. (2000), *Equilibrium Unemployment Theory*, 2nd edition, MIT Press, Cambridge.

[40] POSTEL-VINAY, FABIAN AND JEAN MARC ROBIN (2002), Equilibrium wage dispersion with worker and employer heterogeneity, *Econometrica*, 70, 2295–2350.

- [41] ROSEN, SHERWIN (1974), Hedonic prices and implicit markets: Product differentiation in pure competition, *Journal of Political Economy*, 82, 34-55.
- [42] SATTINGER, MICHAEL (1975), Comparative advantage and the distribution of earnings and abilities, *Econometrica*, 43, 455-68.
- [43] SHIMER, ROBERT (2005), The cyclical behavior of equilibrium unemployment and vacancies, *American Economic Review*, 95, 25-49.
- [44] SHIMER, ROBERT (2006), On-the-job search and strategic bargaining, *in*: H. Bunzel, B.J. Christensen, G.R. Neumann and J.M. Robin, eds., *Structural Models of Wage and Employment Dynamics*, conference volume in honor of Dale Mortensen, Elsevier, Amsterdam.

- [45] SHIMER, ROBERT (2007), Reassessing the ins and outs of unemployment, NBER working paper 13421.
- [46] SHIMER, ROBERT AND LONES SMITH (2000), Assortative matching and search, *Econometrica*, 68, 343-369.
- [47] TEULINGS, C.N. (1995), The wage distribution in a model of the assignment of skills to jobs, *Journal of Political Economy*, 103, 280-315.
- [48] TEULINGS, COEN N. (2005), Comparative advantage, relative wages, and the accumulation of human capital, *Journal of Political Economy*, 113, 425-461.
- [49] TEULINGS, COEN N. AND PIETER A. GAUTIER (2004), The right man for the job, *Review of Economic Studies*, 71, 553-580.

[50] TEULINGS, COEN N. AND THIJS VAN RENS (2008), Education, growth, and income

inequality, *Review of Economics and Statistics*, 90, 89-104.

A Appendix

A.1 Reducing the number of parameters

Consider the economy described in Section 2.2. Take the expressions for $W(x)$ and \bar{x}

from Gautier, Teulings, and Van Vuuren (2010) for the no commitment case (equation

20), using $Sv = 1$ ²¹ and $\kappa \equiv \lambda/(\rho + \delta)$ yields,

$$W(x) = 1 - \gamma \left[\frac{1 + \psi\kappa x}{(\psi\kappa)^2} \log \left(\frac{1 + \psi\kappa x}{1 + \psi\kappa \bar{x}} \right) - \frac{x - \bar{x}}{\psi\kappa} - \frac{1}{2}x(x - 2\bar{x}) \right].$$

$G(x)$ is given by

$$G(x) = 1 - \frac{\bar{x} - x}{(1 + \psi\kappa x)\bar{x}}. \quad (21)$$

²¹In their case, the model is even more extended, $\lambda = \lambda_0 Sv$. We just substituted λ for this combination.

where \bar{x} follows from (1).

A.2 $\frac{f_{ee}}{f_{eu}}$ ratio

$$\begin{aligned} f_{ee} &= (\rho + \delta) \frac{\int_0^{\bar{x}} \psi(1-u) g(x) x dx}{(1-u)} = (\rho + \delta) \psi \frac{(1 + \psi \bar{x})}{\psi^2 \bar{x}} \int_0^{\psi \bar{x}} \frac{x}{(1+x)^2} dx \\ &= (\rho + \delta) \left[\frac{1 + \psi \bar{x}}{\psi \bar{x}} \ln(\psi \bar{x} + 1) - 1 \right]. \end{aligned}$$

using $f_{eu} = \rho + \delta$ yields equation (20).

A.3 The proof of Proposition 2

Let $x \sim \mathcal{N}(0, \sigma_x^2)$ and $\varepsilon_x \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. x and ε_x are independent. Let $w \equiv \ln W - \mathbb{E}[\ln W]$.

The data-generating process for w is,

$$w = \omega_0 - \omega_2 |x|$$

where $\omega_0 = \omega_2 \mathbb{E}|x|$.

The regression we run is

$$w = \bar{\omega}_0 + \bar{\omega}_2(x + \varepsilon_x)^2 + \nu.$$

The claim is

$$\bar{\omega}_0 \rightarrow \frac{1}{2}R\omega_2\mathbb{E}|x|,$$

where $R = \frac{\text{Var}[x]}{\text{Var}[x] + \text{Var}[\varepsilon_x]}.$

CLAIM 1: $\mathbb{E}|x| = \sqrt{\frac{2}{\pi}}\text{Var}[x]$ and $\omega_0 = \omega_2\sqrt{\frac{2}{\pi}}\text{Var}[x].$

Proof.

$$\mathbb{E}|x| = \text{Var}[x] \int_0^\infty x \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

Let $s = x^2/2$, then we have $x = \sqrt{2s}$ and $dx = \frac{1}{\sqrt{2s}}ds$.

$$\begin{aligned}\mathbb{E}|x| &= \text{Var}[x] \sqrt{\frac{2}{\pi}} \int_0^\infty e^{-s} ds \\ &= \text{Var}[x] \sqrt{\frac{2}{\pi}}\end{aligned}$$

Note we have $\omega_0 = \omega_2 \mathbb{E}|x|$. ■

CLAIM 2: $\bar{\omega}_0 = -\bar{\omega}_2(\text{Var}[x] + \text{Var}[\varepsilon_x])$

Proof.

$$\begin{aligned}0 = w &= \bar{\omega}_0 + \bar{\omega}_2 \mathbb{E}(x + \varepsilon_x)^2 \\ &= \bar{\omega}_0 + \bar{\omega}_2(\text{Var}[x] + \text{Var}[\varepsilon_x])\end{aligned}$$

Note $\bar{\omega}_2 \rightarrow \frac{\text{cov}[(x+\varepsilon_x)^2, \omega_0 - \omega_2|x|]}{\text{var}[(x+\varepsilon_x)^2]}$. ■

CLAIM 3: $\text{Var}[(x + \varepsilon_x)^2] = 2(\text{Var}[x] + \text{Var}[\varepsilon_x])^2$.

Proof. Note that $\mathbb{E}x^4 = 3\text{Var}[x]^2$, and x is independent of ε_x . ■

CLAIM 4: $\text{cov}[(x + \varepsilon_x)^2, \omega_0 - \omega_2|x|] = -\text{Var}[x]^{3/2} \cdot \sqrt{\frac{2}{\pi}} \omega_2$.

Proof. Since $\mathbb{E}(\omega_0 - \omega_2|x|) = 0$,

$$\begin{aligned} \text{cov}[(x + \varepsilon_x)^2, \omega_0 - \omega_2|x|] &= \mathbb{E}[(x^2 + 2x\varepsilon_x + \varepsilon_x^2)(\omega_0 - \omega_2|x|)] \\ &= \mathbb{E}(\omega_0 x^2 - \omega_2|x|) \\ &= \text{Var}[x]^{3/2} \omega_0 \int_0^\infty x^2 \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &\quad - \text{Var}[x]^{3/2} \omega_2 \int_0^\infty x^3 \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \end{aligned}$$

Now we can use the same transformation as we did previously. Let $s = x^2$.

$$\begin{aligned} \mathbb{E}(\omega_0 x^2 - \omega_2|x|^3) &= \sigma_x^2 \omega_0 \frac{2}{\sqrt{\pi}} \int_0^\infty s^{1/2} e^{-s} ds - \text{Var}[x]^{3/2} \omega_2 \frac{2\sqrt{2}}{\sqrt{\pi}} \int_0^\infty s e^{-s} ds \\ &= \text{Var}[x] \omega_0 \frac{2}{\sqrt{\pi}} \Gamma\left(\frac{3}{2}\right) - \text{Var}[x]^{3/2} \omega_2 \frac{2\sqrt{2}}{\sqrt{\pi}} \Gamma(2) \\ &= -\text{Var}[x]^{3/2} \sqrt{\frac{2}{\pi}} \omega_2. \end{aligned}$$

where Γ is the gamma function and in the last step we use claim 1, $\omega_0 = \omega_2 \sqrt{\frac{2}{\pi}} \text{Var}[x]^{3/2}$.

■

Proposition 2 $\bar{\omega}_0 \rightarrow \frac{1}{2} R \omega_2 \mathbb{E}|x| = \frac{1}{2} R \omega_0$.

Proof.

$$\begin{aligned}
\bar{\omega}_0 &= -\bar{\omega}_2 (\text{Var}[x] + \text{Var}[\varepsilon_x]) \\
&= -\frac{\text{Cov}[(x + \varepsilon_x)^2, \omega_0 - \omega_2 |x|]}{\text{Var}[(x + \varepsilon_x)^2]} (\text{Var}[x] + \text{Var}[\varepsilon_x]) \\
&= \frac{\omega_2}{\sqrt{2\pi}} \frac{\text{Var}[x]^{3/2}}{\text{Var}[x] + \text{Var}[\varepsilon_x]} \\
&= \frac{1}{2} R \omega_2 \mathbb{E}|x|
\end{aligned}$$

where in the third step, we use claim 3. ■

B Robustness checks

B	γ	u (%)		ψ		$W(0) - E_x W$ x100		$\frac{E_x W}{W(\bar{x})}$		$\frac{1}{2}\gamma\sigma_x^2$ (%)		R (%)	
commitment		yes	no	yes	no	yes	no	yes	no	yes	no	yes	no
0.4	0.6	3.95	3.62	0.26	0.24	4.80	5.72	1.20	1.17	2.01	2.22	56	62
0.4	1.8	6.38	5.99	0.43	0.40	7.31	8.46	1.35	1.29	3.06	3.29	28	31
0.6	0.6	6.28	6.11	0.42	0.41	4.80	5.72	1.20	1.17	2.01	2.22	56	62
0.6	1.8	10.53	11.01	0.74	0.78	7.31	8.01	1.35	1.44	3.06	3.29	28	31

Table 3: Calibration results for $fee/feu = 1.3$

B	γ	u (%)		ψ		$W(0) - E_x W$ x100		$\frac{E_x W}{W(\bar{x})}$		$\frac{1}{2}\gamma\sigma_x^2$ (%)		R (%)	
commitment		yes	no	yes	no	yes	no	yes	no	yes	no	yes	no
0.4	0.6	3.08	2.84	0.37	0.34	4.09	5.53	1.29	1.25	1.57	1.65	44	46
0.4	1.8	4.72	4.48	0.58	0.54	6.04	8.00	1.51	1.44	2.32	2.39	22	22
0.6	0.6	4.82	4.69	0.59	0.57	4.09	5.53	1.29	1.25	1.57	1.65	44	46
0.6	1.8	7.60	7.85	0.95	0.99	6.05	8.01	1.51	1.44	2.32	2.39	22	22

Table 4: Calibration results for $fee/feu = 1.75$

γ	0.6				1.8			
B	0.4		0.6		0.4		0.6	
Commitment	yes	no	yes	no	yes	no	yes	no
$u(1 - B)$	2.37	2.17	2.51	2.44	3.83	3.60	4.21	4.41
vK	3.86	8.38	3.77	8.16	5.73	12.09	5.47	11.44
$(1 - u) [Y(0) - E_x Y]$	1.93	2.14	1.88	2.09	2.86	3.09	2.74	2.93
X (%)	8.16	12.69	8.16	12.69	12.42	18.78	12.42	18.78

Table 5: Decomposition of output loss due to frictions for $f_{ee}/f_{eu} = 1.3$

γ	0.6				1.8			
B	0.4		0.6		0.4		0.6	
Commitment	yes	no	yes	no	yes	no	yes	no
$u(1 - B)$	1.85	1.71	1.93	1.88	2.84	2.69	3.04	3.14
vK	3.05	7.36	2.99	7.22	4.43	10.48	4.30	10.11
$(1 - u) [Y(0) - E_x Y]$	1.52	1.60	1.50	1.57	2.22	2.28	2.15	2.20
X (%)	6.42	10.67	6.42	10.67	9.48	15.45	9.48	15.45

Table 6: Decomposition of output loss due to frictions for $f_{ee}/f_{eu} = 1.75$

C Web Appendix

C.1 Flow conditions

A worker accepts any job offer with a wage above his current wage and consequently with a mismatch indicator smaller than in his current job. Unemployed workers accept only job offers with $x < \bar{x}$. The unemployment rate in this economy is determined by the following equilibrium flow condition

$$\delta(1 - u) + \rho = 2\lambda\bar{x}u + \rho u. \quad (22)$$

The left hand side of (22) measures the inflow into unemployment. The first term reflects workers who loose their job and the second term reflects the growth of the labor force (new workers start as unemployed). The right hand side measures the outflow. The first term is the number of workers who find a job and the second term captures the fact

that the inflow should exceed the outflow by ρu to keep unemployment at a constant fraction of the workforce at the balanced growth (at rate ρ) path. So for $\rho \rightarrow 0$ this is just a simple steady-state flow equation. This relation can be simplified by defining the parameter $\kappa \equiv 2\lambda/(\rho + \delta)$. Then,

$$u = \frac{1}{1 + \kappa \bar{x}}.$$

Let $G[W(x)] \equiv 1 - \hat{G}(x)$. Equation (8) follows from substituting (22) into the following balanced growth equation,

$$2\lambda x \{u + \psi(1 - u)[1 - G(x)]\} - \delta(1 - u)G(x) = \rho(1 - u)G(x).$$

which tells us that the labor force grows at rate ρ and so does the mass of workers who are employed at a distance x or less from their optimal job type. The first term on the left is the inflow into this class and the second term is the outflow. Flows from one

smaller-than- x job to another cancel out. Equation (8) follows from substituting (22) into

the following balanced growth equation,

$$2\lambda x \{u + \psi(1 - u) [1 - G(x)]\} - \delta(1 - u)G(x) = \rho(1 - u)G(x).$$

which tells us that the labor force grows at rate ρ and so does the mass of workers

who are employed at a distance x or less from their optimal job type. The first term on

the left is the inflow into this class and the second term is the outflow. Flows from one

smaller-than- x job to another cancel out.

C.2 Derivation of the Bellman equations under golden growth

A general way to derive V^U and the free entry condition uses a simple accounting identity.

First we start with a general discount rate r , and then we let r approach the population

growth rate ρ from above.

At time 0,

$$\begin{aligned}
uV^U + (1-u)\mathbb{E}V^E + \int_0^\infty V^U e^{-rt} de^{\rho t} &= uV^U + (1-u)\mathbb{E}V^E + \int_0^\infty e^{-rt} V^U \rho e^{\rho t} dt \\
&= \int_0^\infty e^{-rt} e^{\rho t} (uB + (1-u)\mathbb{E}W) dt
\end{aligned}$$

The first term on the first line is the total discounted value created by unemployed workers

at time 0; the second term is the total discounted value created by employed workers at

time 0; the third term is the total discounted value of all future generations, where each

new worker starts his career as an unemployed worker. The second line is an alternative

way to aggregate total value in the economy that equals the left hand side because workers

are risk neutral. It simply equals the discounted total value (expected wage income + B)

of all workers from $t = 0$ onwards. The above equality is essentially an application of the

Fubini theorem.

Simplifying the above equation yields,

$$\begin{aligned}
& uV^U + (1-u)\mathbb{E}V^E + \rho V^U \left[\frac{1}{\rho-r} e^{(\rho-r)t} \right]_0^\infty \\
&= (uB + (1-u)\mathbb{E}W) \left[\frac{1}{\rho-r} e^{(\rho-r)t} \right]_0^\infty
\end{aligned}$$

$$V^U + (r-\rho)(uV^U + (1-u)\mathbb{E}V^E) = uB + (1-u)\mathbb{E}W$$

Letting $r \searrow \rho$ gives

$$\rho V^U = uB + (1-u)\mathbb{E}W.$$

The derivation for the free entry condition is similar. Suppose that at time 0 the population is 1 and the number of vacancies equals v . Denote the value of a filled job by

V^J . Again, there exists an accounting identity for the total value created by firms in the

economy,

$$(1-u)\mathbb{E}V^J + vV^V + \int_0^\infty e^{-rt} V_V \hat{\rho} e^{\rho t} dt = \int_0^\infty e^{-rt} e^{\rho t} ((1-u)(\mathbb{E}Y - \mathbb{E}W) - Kv)$$

where $\hat{\rho}$ is the adjusted birth rate for vacancies (which is not important because in equilibrium $V_V = 0$). The expected value of the cross section of filled and vacant jobs should equal the discounted sum of firm profits minus the amount of resources spent on vacancies.

Since under free entry the expected value of all current and future vacancies equals 0, we get,

$$(r - \rho)(1 - u)\mathbb{E}V^J = (1 - u)(\mathbb{E}Y - \mathbb{E}W) - Kv$$

Letting $r \searrow \rho$ gives

$$vK = (1 - u)(\mathbb{E}Y - \mathbb{E}W).$$

The derivation of $V^E(x)$ follows Gautier et al. (2010). The Bellman equation for the

asset value of employment for a worker employed in a job with mismatch indicator x reads

$$\rho V^E(x) = \widehat{W}(x) + 2\psi\lambda \int_0^x [V^E(y) - V^E(x)] dy - \delta [V^E(x) - V^U]. \quad (23)$$

Totally differentiating yields

$$(\rho + \delta) V_x^E(x) = \frac{\widehat{W}_x(x)}{1 + \psi x}.$$

The solution to this differential equation reads

$$(\rho + \delta) V^E(x) = \int_0^x \frac{\widehat{W}_x(y)}{1 + \psi y} dy + C_0.$$

Integrating by parts yields

$$(\rho + \delta) V^E(x) = \frac{\widehat{W}(x)}{1 + \psi x} - \widehat{W}(0) + \int_0^x \frac{\psi \widehat{W}(y)}{(1 + \psi y)^2} dy + C_0.$$

Evaluation of this equation at $x = 0$ gives an initial condition that can be used to solve

for C_0

$$C_0 = \widehat{W}(0) + \delta V^U.$$

Substitution of this initial condition yields

$$(\rho + \delta) V^E(x) = \frac{\widehat{W}(x)}{1 + \psi x} + \int_0^x \frac{\psi \widehat{W}(y)}{(1 + \psi y)^2} dy + \delta V^U. \quad (24)$$

Define $E_x W = \int_0^{\bar{x}} \widehat{W}(x) d\widehat{G}(x)$. By equation (8), we have

$$\int_0^x \frac{\psi \widehat{W}(y)}{(1 + \psi y)^2} dy = \frac{\psi \bar{x}}{1 + \psi \bar{x}} E_x W.$$

Using this expression and $V^U = V^E(\bar{x})$ to evaluate equation (23) at \bar{x} yields

$$\rho V^E(\bar{x}) = \frac{\widehat{W}(\bar{x}) + \psi \bar{x} E_x W}{1 + \psi \bar{x}} = \frac{u \widehat{W}(\bar{x}) + \psi (1 - u) E_x W}{u + \psi (1 - u)}. \quad (25)$$

Equation (23) and $V^U = V^E(\bar{x})$ imply

$$2\psi \lambda \int_0^{\bar{x}} [V^E(x) - V^U] dx = \rho V^U(x) - \widehat{W}(\bar{x}).$$

Substitution of this expression into the Bellman equation for ρV^U yields

$$\rho V^U = B + 2\lambda \int_0^{\bar{x}} [V^E(x) - V^U] dx = \frac{B + \bar{x} E_x W}{1 + \bar{x}} = uB + (1 - u) E_x W, \quad (26)$$

where the final step uses $u = 1/(1 + \bar{x})$.

Finally, consider the free entry condition. Define $E_G Y \equiv \int_0^{\bar{x}} g(x) Y(x) dx$ and $E_G W \equiv$

$\int_0^{\bar{x}} g(x) W(x) dx$, then,

$$\begin{aligned} K &= 2\lambda \int_0^{\bar{x}} \{u + \psi(1 - u)[1 - G(x)]\} \frac{Y(x) - W(x)}{\rho + \delta + 2\psi\lambda vx} dx \\ &= \frac{1 - u}{v} (E_G Y - E_G W). \end{aligned} \quad (27)$$

The term, $u + \psi(1 - u)[1 - G(x)]$ is the effective labor supply for a type x match. The

second term gives the value of a filled vacancy. It discounts current revenue $Y(x) - W(x)$

by the discount rate ρ plus the separation rate δ plus the quit rate $2\psi\lambda x$. The final line

follows from substituting in (8) and $(\bar{x} = (1 - u)/u)$. This implies that the resources spent

on vacancy creation, vK , must in a steady-state equilibrium be equal to the employment

rate $(1 - u)$ times the expected profit of a filled vacancy, $(E_G Y - E_G W)$.

C.3 Derivation of wages

This derivation summarizes the results in Gautier et al. (2010). Conditional on x , firms

choose a wage that maximizes the value of a vacancy,

$$\arg \max_W \left(\left\{ u + \psi (1 - u) \left[1 - \widehat{G}(W) \right] \right\} \frac{Y(x) - W}{\psi \kappa \bar{x} \widehat{F}(W)} \right)$$

The FOC with respect to W reads

$$0 = - \frac{\psi (1 - u) G_x / W_x}{u + \psi (1 - u) \left[1 - \widehat{G}(W) \right]} - \frac{\psi \kappa \bar{x} F_x / W_x}{\psi \lambda \bar{x} \widehat{F}(W)} - \frac{1}{Y(x) - W},$$

where we use $\widehat{G}_W = G_x/W_x$ and $\widehat{F}_W = F_x/W_x$. Using $F_x = 1/\bar{x}$ and some rearrangement

yields,

$$W_x = - \left(\frac{\psi (1-u) G_x}{u + \psi (1-u) [1 - \widehat{G}(W)]} + \frac{\psi \kappa}{1 + \psi \kappa x} \right) [Y(x) - W]. \quad (28)$$

Use (8) and its derivative with respect to x to write,

$$\frac{\psi (1-u) G_x}{u + \psi (1-u) [1 - \widehat{G}(W)]} = \frac{\psi \kappa}{1 + \psi \kappa x},$$

and substitute this back in (28) to get

$$W_x(x) = - \frac{2\psi \kappa (Y(x) - W(x))}{1 + \psi \kappa x}.$$

This equation is almost identical to equation to the one for no commitment except that

the "2" is replaced by a "1" (at the margin the hiring and no quit premia are equal and

under no commitment there are no hiring premia). For the solution of the differential

equation we refer to Gautier et al. (2010).■

C.4 Wages and expected wages

Here we combine γ and κ into one parameter ζ (as we do in our matlab program). Define,

$\gamma \equiv \kappa^2 \zeta$, $x \equiv z/\kappa$, and $\bar{x} = \bar{z}/\kappa$. Note that we can switch back and forth between the

model in terms of (x, γ) and (z, ζ) by the fact that $\gamma \text{Var}[x] = \zeta \text{Var}[z]$. $G(x)$ is given by,

$$G(x) = 1 - \frac{\bar{x} - x}{(1 + \psi \kappa x) \bar{x}}.$$

Substitution of the expressions for x and \bar{x} in (21) shows that $\hat{G}(z)$ satisfies

$$\hat{G}(z) = 1 - \frac{\bar{z} - z}{(1 + \psi z) \bar{z}}.$$

The value of \bar{z} follows from (7), using $1 - u = \bar{z}/(1 + \bar{z})$ and $V^E(\bar{x}) = V^U$. Substituting

those variables in the wage equations (18,19) of Gautier, Teulings yields,

commitment

$$\widehat{W}(z) = 1 - \zeta \left[\left(\frac{1 + \psi z}{\psi} \right)^2 \log \left(\frac{1 + \psi \bar{z}}{1 + \psi z} \right) - \frac{\bar{z}}{\psi} \frac{(1 + \psi z)^2}{1 + \psi \bar{z}} + \frac{z}{\psi} + \frac{3}{2} z^2 \right], \quad (29)$$

$$\begin{aligned} \mathbb{E}_z W &= \int_0^{\bar{z}} g(z) \widehat{W}(z) dz \\ &= 1 - \int_0^{\bar{z}} \frac{1 + \psi \bar{z}}{\bar{z}(1 + \psi z)^2} \zeta \left[- \left(\frac{1 + \psi z}{\psi} \right)^2 \ln \left(\frac{1 + \psi z}{1 + \psi \bar{z}} \right) - \frac{\bar{z}}{\psi} \frac{(1 + \psi z)^2}{1 + \psi \bar{z}} + \frac{z}{\psi} + \frac{3}{2} z^2 \right] dz \\ &= 1 - \zeta \frac{1 + \psi \bar{z}}{\psi^3 \bar{z}} \int_0^{\psi \bar{z}} \frac{1}{(1 + q)^2} \left[- (1 + q)^2 \ln \left(\frac{1 + q}{1 + \psi \bar{z}} \right) - \psi \bar{z} \frac{(1 + q)^2}{1 + \psi \bar{z}} + q + \frac{3}{2} q^2 \right] dq \\ &= 1 - \zeta \frac{3}{\psi^3 \bar{z}} \left[\psi \bar{z} + \frac{1}{2} \psi^2 \bar{z}^2 - (1 + \psi \bar{z}) \ln(1 + \psi \bar{z}) \right]. \end{aligned}$$

The expression for ζ can be derived from combining $\bar{z} = \frac{1-u}{u}$, (6) and (7)

$$\zeta = \frac{2\psi^3 (1 - B) (1 + \bar{z})(1 + \psi \bar{z})}{6(1 - \psi)(1 + \psi \bar{z})(1 + \bar{z}) \ln(1 + \psi \bar{z}) - \psi \bar{z} [6(1 + \bar{z}) + 3\psi(2 - \bar{z})(1 + \bar{z}) + \psi^2 \bar{z}(4 + 3\bar{z})]}.$$

non-commitment

$$\widehat{W}(z) = 1 - \zeta \left[\frac{1 + \psi z}{\psi^2} \ln \left(\frac{1 + \psi z}{1 + \psi \bar{z}} \right) - \frac{z - \bar{z}}{\psi} - \frac{1}{2} z (z - 2\bar{z}) \right], \quad (30)$$

$$\begin{aligned} \mathbb{E}_z W &= \int_0^{\bar{z}} g(z) \widehat{W}(z) dz \\ &= 1 - \int_0^{\bar{z}} \frac{1 + \psi \bar{z}}{\bar{z} (1 + \psi z)^2} \zeta \left[\frac{1 + \psi z}{\psi^2} \ln \left(\frac{1 + \psi z}{1 + \psi \bar{z}} \right) - \frac{z - \bar{z}}{\psi} - \frac{1}{2} z (z - 2\bar{z}) \right] dz \\ &= 1 - \zeta \frac{1 + \psi \bar{z}}{\psi^3 \bar{z}} \int_0^{\psi \bar{z}} \frac{1}{(1 + q)^2} \left[(1 + q) \ln \left(\frac{1 + q}{1 + \psi \bar{z}} \right) - q + \psi \bar{z} - \frac{1}{2} q (q - 2\psi \bar{z}) \right] dq \\ &= 1 - \zeta \frac{1 + \psi \bar{z}}{\psi^3 \bar{z}} \left[-\frac{1}{2} \ln^2 (1 + \psi \bar{z}) + \psi \bar{z} \ln (1 + \psi \bar{z}) - \frac{1}{2} \frac{\psi^2 \bar{z}^2}{1 + \psi \bar{z}} \right], \end{aligned}$$

Again, ζ can be derived from $\bar{z} = \frac{1-u}{u}$, (6) and (7),

$$\zeta = \frac{2\psi^3 (1 - B) (1 + \psi \bar{z}) (1 + \bar{z})}{\psi^2 (1 + \psi \bar{z}) (1 + \bar{z}) \bar{z}^2 + \bar{z}^2 - (1 - \psi) (1 + \bar{z})^2 [2\psi \bar{z} - \ln (1 + \psi \bar{z})] \ln (1 + \psi \bar{z})}.$$

C.5 Variance of x

$$\begin{aligned}\text{Var}[x] &= \int_0^{\bar{x}} x^2 g(x) dx = \frac{1 + \psi\bar{x}}{\bar{x}} \int_0^{\bar{x}} \frac{x^2}{(1 + \psi x)^2} dx = \frac{1 + \psi\bar{x}}{\psi^3 \bar{x}} \int_0^{\psi\bar{x}} \frac{x^2}{(1 + x)^2} dx \\ &= \frac{\psi\bar{x}(2 + \psi\bar{x}) - 2(1 + \psi\bar{x})\log(1 + \psi\bar{x})}{\psi^3 \bar{x}}.\end{aligned}\quad (31)$$

C.6 Measuring γ

Let $\tilde{Y}(s, c)$ be the productivity of an s -type worker in a c -type job. We can adjust the

production function to be increasing in s , as follows,

$$\ln \tilde{Y}(s, c) = s - \frac{1}{2}\gamma(s - c)^2 = s - \frac{1}{2}\gamma x^2. \quad (32)$$

The optimal assignment $c(s)$ for worker type s solves the first-order condition $\ln \tilde{Y}_c[s, c(s)] =$

0, implying that $c(s) = s$. This specification exhibits all features discussed in Sec-

tion 3.1. First, in the Walrasian equilibrium, all workers are assigned to their opti-

mal job type, $c(s) = s$ (and hence $x = 0$) and wages are equal to output. Hence

$\ln \tilde{Y}(s, s) = \ln \widehat{W}(s) = s$, see equation (10). Second, the function $\tilde{Y}(s, c)$ exhibits

absolute and comparative advantage which is required for the derivation of \hat{s} and \hat{c} as

discussed in Section (3.1).²² Third, the Mincerian rate of return to skill $dE_x[\ln W]/ds$ is

equal to one. Finally, the parameter γ in equation (32) is equivalent to the parameter γ

in equation (1) by applying a second-order Taylor expansion to $Y(x)$ and $\tilde{Y}(s, c)$, around

the optimal assignment, $x = 0$: $\gamma = Y''(0)/Y(0) = \tilde{Y}_{cc}(c, c)/\tilde{Y}(c, c)$. The parameter γ

measures the curvature of Rosen's (1974) well known hedonic system where the isoprofit

curve and the indifference curve of the worker are tangent to the locus of market wages

and the indifference curve of the worker. This curvature, the second derivative of the

production function, is a measure of both the productivity loss due to mismatch and the

²²Comparative advantage requires the cross derivative of $\ln Y(s, c)$ to be positive; absolute advantage requires the first derivative of $\ln Y(s, c)$ with respect to s to be positive for any c . The latter is not globally satisfied for this polynomial specification, but it is in the optimal assignment $s = c$.

elasticity of substitution between various workers types. Note that the dimension of this

curvature parameter γ corresponds to that of the mismatch indicator x used in Section

(3.1) since the Mincerian rate of return to s is equal to unity in the optimal assignment:

$$\ln \tilde{Y}_s(s, s) = 1.$$

C.7 Constructing skill and job complexity levels

We run the following two regressions

$$\ln W = \beta \vec{j} + \varepsilon_s + \varepsilon_w, \tag{33}$$

$$\ln W = \alpha \vec{k} + \varepsilon_c + \varepsilon_w,$$

where \vec{j} and \vec{k} are vectors of observed worker and job characteristics respectively, where

ε_s and ε_c capture unobserved worker and job characteristics respectively, and where ε_w

captures the effect of non-optimal assignment on wages and measurement error in wages.

It is convenient to normalize our data on $\ln W$, \vec{j} and \vec{k} such that they have zero mean.

Hence, it does not make sense to include a constant in these regressions. The estimated

parameter vector can then be used to construct indices for the observed worker and job

characteristics $\hat{s} \equiv \beta \vec{j}$, and $\hat{c} \equiv \alpha \vec{k}$. The non-linearities in the relation between $\ln W$ on

the one hand and \vec{j} and \vec{k} on the other are captured by a proper normalization. Then,

the skill index s and the job index c satisfy,

$$s = \hat{s} + \varepsilon_s,$$

$$c = \hat{c} + \varepsilon_c,$$

Both indices \hat{s} and \hat{c} have zero mean by construction and are uncorrelated to the unob-

served components ε_s and ε_c , respectively.²³

²³We apply the following iterative procedure such that if we regress $\ln W$ on both \bar{s} and \bar{s}^2 that the coefficient of the second-order term \bar{s}^2 is zero. First, run $\ln W = \chi_1 \hat{s}_1 + \chi_2 \hat{s}_1^2 + \underline{\varepsilon}_{s1}$, where \hat{s}_1 is $E(s|\vec{j})$

constructed from (33) and where $\varepsilon_{s_1} = s - \hat{s}_1$. Second, we construct a new variable $\hat{s}_2 = \chi_1 \hat{s}_1 + \chi_2 \hat{s}_1^2 - \text{E}[\chi_1 \hat{s}_1 + \chi_2 \hat{s}_1^2]$ and rerun the first regression after substituting \bar{s}_2 for \bar{s}_1 . We repeat these steps until $\chi_2 = 0$. The same applies to our regression for \bar{c} . This algorithm therefore normalizes \bar{s} in such a way that any correlation between \hat{s} and ε_s is eliminated.